Aalto University
School of Science
Degree Programme in Engineering Physics and Mathematics

Jaakko Hujanen

Inverse Problem for Quantum Mechanical Scattering in One Dimension

Master's Thesis
Espoo, August 28, 2014

Supervisor: Professor Rolf Stenberg
Advisors: Associate Professor Nuutti Hyvönen
D.Sc. (Tech.) Teemu Ojanen

Aalto University
School of Science
Degree Programme in Engineering Physics and Mathematics

ABSTRACT OF
MASTER'S THESIS

Author: Jaakko Hujanen

Title:
Inverse Problem for Quantum Mechanical Scattering in One Dimension

Date: August 28, 2014 **Pages:** 69

Major: Mechanics **Code:** Mat-5

Supervisor: Professor Rolf Stenberg

Advisors: Associate Professor Nuutti Hyvönen
D.Sc. (Tech.) Teemu Ojanen

The field of inverse problem considers mathematical methods for reconstructing unknown parameters of some model from given measurements. Such problems arise in different branches of science, e.g, in medical imaging. In this thesis, inverse mathematics is applied to a one-dimensional scattering problem in a quantum mechanical setting: the aim is to create an algorithm to determine the potential function of a scattering interval when transmission of electrons is observed. Since the scattering problem includes non-linear phenomena, an iterative reconstruction scheme is employed. The implemented numerical method combines finite element method (FEM) and the Gauss–Newton optimization algorithm in a manner which utilizes the FEM solver for numerical evaluations of the Jacobian matrix. Furthermore, the forward problem is shown to have a unique solution to reinforce credibility of the procedure. The resulting program optimally gives good reconstructions with simulated test data, but it is sensitive to the choice of regularization.

Keywords: inverse problems, quantum mechanical scattering, finite element method, Gauss–Newton method

Language: English

Aalto-yliopisto
Perustieteiden korkeakoulu
Teknillisen fysiikan ja matematiikan koulutusohjelma

DIPLOMITYÖN TIIVISTELMÄ

Tekijä:	Jaakko Hujanen		
Työn nimi:	Yksiulotteisen kvanttimekaanisen sironnan käänteisongelma		
Päiväys:	28. elokuuta 2014	**Sivumäärä:**	69
Pääaine:	Mekaniikka	**Koodi:**	Mat-5
Valvoja:	Professori Rolf Stenberg		
Ohjaajat:	Professori Nuutti Hyvönen TkT Teemu Ojanen		

Käänteisongelmissa tavoitteena on päätellä fysikaalisen systeemin parametreja mittausdatasta. Tällaisia ongelmia esiintyy tieteessä laajalti, ja sovelluskohteita on esimerkiksi lääketieteellisessä kuvantamisessa. Tämän työn tavoitteena on luoda algoritmi kvanttimekaaniselle sirontaongelmalle, jossa tutkittavan alueen potentiaalifunktion ominaisuuksia selvitetään havaituista elektronien läpäisytodennäköisyyksistä. Tässä tapauksessa sironnan matemaattinen malli on epälineaarinen, minkä takia joudutaan käyttämään elementtimenetelmään (FEM) perustuvaa iteratiivista numeerista inversioalgoritmia. Ratkaisin pohjautuu Gauss–Newton-optimointimenetelmään, jossa Jacobin matriisi lasketaan numeerisesti FEM-ratkaisijalla. Suoran sirontaongelman ratkaisun olemassaolo ja yksikäsitteisyys todistetaan algoritmin toimivuuden varmistamiseksi. Potentiaalien rekonstruktio testidatasta onnistuu hyvin, mutta menetelmä on herkkä regularisaation valinnan suhteen.

Asiasanat:	käänteisongelmat, kvanttimekaaninen sironta, elementtimenetelmä, Gauss–Newton-menetelmä
Kieli:	englanti

Acknowledgements

First of all, I would like to sincerely thank my instructors Nuutti Hyvönen and Teemu Ojanen for their great advice, insight and effort which highly contributed to completing the thesis and made the writing process advance smoothly. Likewise, I am grateful for all the help I have received from my colleagues. This has been an interesting project to work on and I am thankful to be given such an opportunity.

Last but not least, I wish to thank all my friends and family for their kind support and pleasant company during free time.

Espoo, August 28, 2014

Jaakko Hujanen

Contents

Chapter 1

Introduction

Quantum mechanics is the theory of microscopic worlds, and along with other modern fields of physics, it has enabled many everyday equipment ranging from computers to USB drives. Understanding many modern applications requires quantum mechanical treatment; consider, for instance, magnetic resonance imaging (MRI) and electron microscopes. The motivation behind the thesis is based on semiconductors which is a large subfield in quantum physics. We are especially interested in the transmission of electrons.

In order to have a solely quantum mechanical system with clear applications, we consider the so-called mesoscopic regime where the object of interest is larger than pure microscopic scale but still not macroscopic which would deteriorate the quantum effects. In particular, for semiconductors the mesoscopic state would mean very low temperatures. One can also observe similar behavior in extremely thin metal wires. Usually, when mesoscopic systems are considered, little importance is put on the object itself; instead, how it alters, e.g., incoming current is investigated. The approach of observing the transformation of inputs to outputs is known as scattering, and in this work it is studied with a simple one-dimensional model.

To introduce a slightly different flavor to the problem, we assume the object has unknown properties which need to be deduced from the scattering information. This sort of procedure is commonly known as an inverse problem, a manner of solving a problem 'backwards': if a model describes a transition from parameters to data, its inverse means deducing model parameters which reproduce the data as well as possible. Inverse problems are widely applied in different fields of science but the most familiar example is probably medical imaging. As an example, in computed tomography, a patient is scanned from multiple directions by x-ray and, for instance, a three-dimensional model of the patient's brain is constructed. In this thesis, we numerically solve an inverse scattering problem in the quantum setting.

This document is organized in three chapters. First, we provide the physical background for our scattering problem and derive the governing equations. In addition, discussion on electron transport in mesoscopic systems is included. All this can be found in Chapter 2. Subsequently, in Chapter 3 we introduce a numerical method for solving a quantum scattering problem in one dimension with a finite element method (FEM). To successfully invert the scattering process, we also need to know how sensitive the forward solution is with respect to the input parameters. The sensitivity is also covered in Chapter 3. Finally, we numerically solve the inverse scattering by utilizing an iterative method. General notes on inverse problems and the derivation of iterative scheme are presented in Chapter 4.

Chapter 2

Physical Background

In this chapter a short introduction to quantum mechanics is given. The basics are covered in Section 2.1 and a scattering problem in one spatial dimension is introduced in Section 2.2. This chapter considers the physical aspects of the thesis and the following ones focus more on the mathematical setup.

2.1 Quantum Mechanics

We start by giving the basic building blocks, postulates, of quantum mechanics and subsequently proceed to the scattering problem. Finally, electron transport is considered as an application of the derived setting. All material in this section is based on the textbook [13].

Postulates

1. Every well-defined physical observable A has a corresponding operator $\hat{A}$. Measured values a of A satisfy the eigenvalue equation

$$\hat{A}\,|\varphi_a\rangle = a\,|\varphi_a\rangle$$

   for some eigenstates $|\varphi_a\rangle$.

2. A measurement of A leaves the system in the eigenstate $|\varphi_a\rangle$.

3. The state of the system at any time is represented by a state vector $|\varphi\rangle$ in a Hilbert space of all states. The expected value of an observable A is obtained as an inner product

$$\mathbb{E}A = \langle\varphi|\,\hat{A}\,|\varphi\rangle$$

when $|\varphi\rangle$ is a unit vector.

4. The time development of the physical system is described by the Schrödinger equation
$$\hat{H}\,|\varphi\rangle = i\hbar\frac{\partial}{\partial t}\,|\varphi\rangle \tag{2.1}$$
where $\hat{H}$ is the Hamiltonian operator.

Discussion

We have described the postulates using the Dirac notation. One could also use a wave function notation. The connection between the two notations originates from the definition of the inner product

$$\langle\varphi|\psi\rangle = \int_{-\infty}^{\infty}\varphi(x,t)^*\psi(x,t)\,dx$$

in a one-dimensional space. Here φ^* denotes the conjugate transpose of φ. Usually the state vector is normalized to unit length,

$$\langle\varphi|\varphi\rangle = 1,$$

which in the wave function form means that

$$\int_{-\infty}^{\infty}\varphi(x,t)^*\varphi(x,t)\,dx = \int_{-\infty}^{\infty}|\varphi(x,t)|^2\,dx = 1.$$

Because of this, we can interpret

$$\int_a^b|\varphi(x,t)|^2\,dx$$

as the probability of finding a particle in the interval $[a,b]$, that is, $\varphi(x,t)$ is the probability density function for finding a particle at x at time t. Thus, the unit normalization means that the particle is somewhere in the space at any given time.

Since the state vectors are elements of a Hilbert space, any superposition of states is also a state. E.g.,

$$|\varphi\rangle = a_1\,|\varphi_1\rangle + a_2\,|\varphi_2\rangle\,.$$

Now, if we measure A, we will get either

$$\hat{A}\,|\varphi\rangle = a_1\,|\varphi_1\rangle$$

or

$$\hat{A}\,|\varphi\rangle = a_2\,|\varphi_2\rangle\,.$$

If $|\varphi\rangle$ is normalized, the probabilities for these events are $|a_1|^2$ and $|a_2|^2$, respectively.

Schrödinger Equation

The Schrödinger equation (2.1) describes the time evolution of a system. There also exists a time-independent version of the equation

$$\hat{H}\,|\varphi\rangle = E\,|\varphi\rangle \tag{2.2}$$

where E is the energy of the state $|\varphi\rangle$. In other words, eigenvalues of the Hamiltonian correspond to observable total energy of the system. In the following, the Hamiltonian of a particle in a one-dimensional potential is written explicitly and the Schrödinger equation for the same setup is presented. The goal is to derive a boundary value problem formulation for our scattering problem. All this is done in the time-independent framework and in the wave function setting.

To begin with, the Hamiltonian is built of two operators: one for the kinetic energy and the other for the potential energy,

$$\hat{H} = \hat{K} + \hat{V}.$$

The kinetic energy operator is

$$\hat{K} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$$

where $\hbar$ is the reduced Planck's constant and m is the mass of the particle, which in our case is an electron. The potential operator is simply

$$\hat{V} = V(x)$$

where the potential V may depend on x. We now write the equation (2.2) as

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + V(x)\psi(x) = E\psi(x).$$

Next magnetic effects are included for electrons which are spin-$\frac{1}{2}$ particles. The potential energy of a magnetic dipole in the classical setting is

$$U = -\boldsymbol{\mu}\cdot\boldsymbol{B},$$

where $\boldsymbol{B}$ is the magnetic field and $\boldsymbol{\mu}$ is the magnetic dipole moment which in quantum mechanics is related to particle spin. Spin is an intrinsic degree of freedom of particles, and it is accommodated in the model by promoting the scalar wave function to a two-component spinor which is an element of $\mathbb{C}^2$:

$$\boldsymbol{\psi} = \begin{bmatrix}\psi^{\uparrow}\\ \psi^{\downarrow}\end{bmatrix} \tag{2.3}$$

where the first component of $\boldsymbol{\psi}$ corresponds to up spin and the second to down spin. In classical physics, magnetic field exerts torque on the dipole to align it with the field. In the quantum setting, this means that the magnetic field and the spin are connected. The operator corresponding to the magnetic potential U is

$$\hat{U} = -\frac{q\hbar}{2m}\boldsymbol{\sigma} \cdot \boldsymbol{B}(x)$$

where the dot operator denotes a linear combination of Pauli matrices $\boldsymbol{\sigma}$, instead of the usual inner product of two vectors, and q is the charge of the particle. The Pauli matrices are

$$\boldsymbol{\sigma} = \left\{ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right\},$$

and thus

$$\boldsymbol{\sigma} \cdot \boldsymbol{B} = B^1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + B^2 \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + B^3 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.4}$$

when the superscripts denote the components of $\boldsymbol{B}$. Our Hamiltonian is then altogether

$$\hat{H} = \hat{K} + \hat{U} + \hat{V}, \tag{2.5}$$

and the whole equation for a particle in a magnetic field becomes

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\boldsymbol{\psi}(x) - \left(\frac{q\hbar}{2m}\boldsymbol{\sigma} \cdot \boldsymbol{B}(x)\right)\boldsymbol{\psi}(x) + V(x)\boldsymbol{\psi}(x) = E\boldsymbol{\psi}(x) \tag{2.6}$$

where $\boldsymbol{\psi}$ is the two-component wave function (2.3).

2.2 Scattering

In this section, we will derive the equations used in the mathematical considerations of Chapter 3. The space has one scattering region, the interval $(0, L)$, and the rest of the space is empty. The setup is shown in Figure 2.1. To find the wave function, we need to solve $\boldsymbol{\psi}(x)$ from

$$\begin{cases} -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\boldsymbol{\psi}(x) - \left(\frac{q\hbar}{2m}\boldsymbol{\sigma} \cdot \boldsymbol{B}(x)\right)\boldsymbol{\psi}(x) + V(x)\boldsymbol{\psi}(x) = E\boldsymbol{\psi}(x), & x \in (0, L), \\ -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\boldsymbol{\psi}(x) = E\boldsymbol{\psi}(x), & x \notin (0, L), \\ \boldsymbol{\psi}(x) \text{ is continuous and continuously differentiable at } 0 \text{ and } L. \end{cases}$$

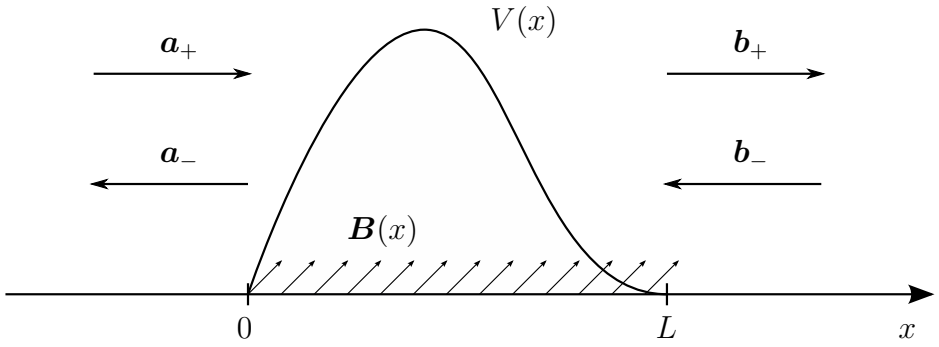

Figure 2.1: Scattering from a potential barrier

In scattering one is naturally not interested in the wave function in the region $(0, L)$. The aim is instead to find the solution in the open space, that is, outside $(0, L)$. It is known that the solution in $\mathbb{R} \setminus [0, L]$ is just a plane wave

$$\boldsymbol{A}_1 e^{ikx} + \boldsymbol{A}_2 e^{-ikx},$$

with

$$k = \frac{\sqrt{2mE}}{\hbar}$$

and some complex coefficients $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$. To differentiate between the solutions in $(-\infty, 0)$ and (L, ∞), we denote

$$\begin{cases} \boldsymbol{\psi}_1(x) = \boldsymbol{a}_+ e^{ikx} + \boldsymbol{a}_- e^{-ikx}, & x \in (-\infty, 0), \\ \boldsymbol{\psi}_2(x) = \boldsymbol{b}_+ e^{ikx} + \boldsymbol{b}_- e^{-ikx}, & x \in (L, \infty). \end{cases}$$

The wave amplitudes $\boldsymbol{a}_\pm$ and $\boldsymbol{b}_\pm$ are drawn in Figure 2.1 as arrows which show the propagation direction of the wave. The entities $\boldsymbol{a}_\pm$ and $\boldsymbol{b}_\pm$ fully describe the solutions in the exterior region, and thus, it is them we want to determine.

We proceed by writing the continuity and differentiability conditions. Letting $\boldsymbol{\psi}$ denote the solution in $(0, L)$, we obtain at $x = 0$ and $x = L$ the identities

$$\begin{cases} \boldsymbol{\psi}(0) = \boldsymbol{\psi}_1(0), \\ \boldsymbol{\psi}'(0) = \boldsymbol{\psi}_1'(0), \\ \boldsymbol{\psi}(L) = \boldsymbol{\psi}_2(L), \\ \boldsymbol{\psi}'(L) = \boldsymbol{\psi}_2'(L). \end{cases}$$

To simplify these conditions, we include L in the differential equation itself instead of the continuity conditions: We scale the spatial variable,

$$\xi = \frac{x}{L},$$

and set

$$\boldsymbol{\varphi}(\xi) = \boldsymbol{\psi}(L\xi).$$

This results in new equations

$$\begin{cases} -\dfrac{\hbar^2}{2mL^2}\dfrac{d^2}{d\xi^2}\boldsymbol{\varphi}(\xi) - \left(\dfrac{q\hbar}{2m}\boldsymbol{\sigma}\cdot\tilde{\boldsymbol{B}}(\xi)\right)\boldsymbol{\varphi}(\xi) + \tilde{V}(\xi)\boldsymbol{\varphi}(\xi) = E\boldsymbol{\varphi}(\xi), & \xi \in (0,1), \\ -\dfrac{\hbar^2}{2mL^2}\dfrac{d^2}{d\xi^2}\boldsymbol{\varphi}(\xi) = E\boldsymbol{\varphi}(\xi), & \xi \notin (0,1), \\ \boldsymbol{\varphi}(\xi) \text{ is continuous and continuously differentiable at 0 and 1,} \end{cases}$$

and the corresponding plane wave solutions are

$$\begin{cases} \boldsymbol{\varphi}_1(\xi) = \boldsymbol{a}_+ e^{i\kappa\xi} + \boldsymbol{a}_- e^{-i\kappa\xi}, & \xi \in (-\infty, 0), \\ \boldsymbol{\varphi}_2(\xi) = \boldsymbol{b}_+ e^{i\kappa\xi} + \boldsymbol{b}_- e^{-i\kappa\xi}, & \xi \in (1, \infty). \end{cases}$$

Here $\tilde{\boldsymbol{B}}$ and $\tilde{V}$ are the scaled magnetic field and potential, respectively, and $\kappa = Lk$. For the scaled equation the continuity and differentiability conditions are

$$\begin{cases} \boldsymbol{\varphi}(0) = \boldsymbol{\varphi}_1(0) = \boldsymbol{a}_+ + \boldsymbol{a}_-, \\ \boldsymbol{\varphi}'(0) = \boldsymbol{\varphi}_1'(0) = i\kappa\boldsymbol{a}_+ - i\kappa\boldsymbol{a}_-, \\ \boldsymbol{\varphi}(1) = \boldsymbol{\varphi}_2(1) = \boldsymbol{b}_+ e^{i\kappa} + \boldsymbol{b}_- e^{-i\kappa}, \\ \boldsymbol{\varphi}'(1) = \boldsymbol{\varphi}_2'(1) = i\kappa\boldsymbol{b}_+ e^{i\kappa} - i\kappa\boldsymbol{b}_- e^{-i\kappa}. \end{cases}$$

These equations describe the boundary conditions for the Schrödinger equation in $(0, 1)$.

We have now all needed information to write the complete set of equations to describe a one-dimensional scattering problem. They are

$$\begin{cases} -\dfrac{\hbar^2}{2mL^2}\dfrac{d^2\boldsymbol{\varphi}(\xi)}{d\xi^2} - \left(\dfrac{q\hbar}{2m}\boldsymbol{\sigma}\cdot\tilde{\boldsymbol{B}}\right)\boldsymbol{\varphi}(\xi) + \tilde{V}\boldsymbol{\varphi}(\xi) = E\boldsymbol{\varphi}(\xi), & \xi \in (0,1), \\ \boldsymbol{\varphi}(0) = \boldsymbol{a}_+ + \boldsymbol{a}_-, \\ \boldsymbol{\varphi}'(0) = i\kappa\boldsymbol{a}_+ - i\kappa\boldsymbol{a}_-, \\ \boldsymbol{\varphi}(1) = \boldsymbol{b}_+ e^{i\kappa} + \boldsymbol{b}_- e^{-i\kappa}, \\ \boldsymbol{\varphi}'(1) = i\kappa\boldsymbol{b}_+ e^{i\kappa} - i\kappa\boldsymbol{b}_- e^{-i\kappa}. \end{cases} \tag{2.7}$$

This boundary value problem, however, cannot be solved without further controlling the unknown coefficients $\boldsymbol{a}_\pm$ and $\boldsymbol{b}_\pm$. In the one-dimensional

case, we are interested in how an incoming wave is transmitted through or reflected back by the scattering region $(0, 1)$. Consequently, without loss of generality, we can assume the incoming wave to be known, that is, $\boldsymbol{a}_+$ is fixed. Also, we may assume that there is no incoming wave from the right, which means that $\boldsymbol{b}_- = \boldsymbol{0}$. These turn out to be sufficient assumptions for guaranteeing the unique solvability of (2.7).

Before we write our final version of the scattering equation, we simplify it by grouping constants and scaling the magnetic field. By moving all wave function terms to the left hand side of the Schrödinger equation and dividing by the coefficient of the second derivate, we get

$$\frac{d^2}{d\xi^2}\boldsymbol{\varphi}(\xi) + \frac{2mL^2}{\hbar^2}\left(E\boldsymbol{\varphi}(\xi) + \left(\frac{q\hbar}{2m}\boldsymbol{\sigma}\cdot\tilde{\boldsymbol{B}}(\xi)\right)\boldsymbol{\varphi}(\xi) - \tilde{V}(\xi)\boldsymbol{\varphi}(\xi)\right) = 0.$$

Defining

$$C = \frac{2mL^2}{\hbar^2},$$

changing the notation back to the original one

$$\xi \hookrightarrow x, \quad \kappa \hookrightarrow k, \quad \tilde{V}(x) \hookrightarrow V(x),$$

and scaling

$$\frac{q\hbar}{2m}\tilde{\boldsymbol{B}}(x) \hookrightarrow \boldsymbol{B}(x),$$

we finally have

$$\frac{d^2}{dx^2}\boldsymbol{\varphi}(x) + C\left(E\boldsymbol{I} + \boldsymbol{\sigma}\cdot\boldsymbol{B}(x) - V(x)\boldsymbol{I}\right)\boldsymbol{\varphi}(x) = 0.$$

Here $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{B}$ now has the same energy units as E and V. Thus $\boldsymbol{B}$ represents the energy effects of the magnetic field instead of the field itself.

Using the assumptions on $\boldsymbol{a}_+$ and $\boldsymbol{b}_-$, we deduce from the constraint equations that

$$\begin{cases} \boldsymbol{\varphi}'(0) + ik\boldsymbol{\varphi}(0) = 2ik\boldsymbol{a}_+, \\ \boldsymbol{\varphi}'(1) - ik\boldsymbol{\varphi}(1) = 0, \end{cases}$$

where

$$k = \frac{L}{\hbar}\sqrt{2mE} = \sqrt{CE}.$$

Finally, the whole boundary value problem becomes

$$\begin{cases} \boldsymbol{\varphi}''(x) + C\left((E - V(x))\boldsymbol{I} + \boldsymbol{\sigma}\cdot\boldsymbol{B}(x)\right)\boldsymbol{\varphi}(x) = 0, \quad x \in (0,1), \\ \boldsymbol{\varphi}'(0) + ik\boldsymbol{\varphi}(0) = 2ik\boldsymbol{a}_+, \\ \boldsymbol{\varphi}'(1) - ik\boldsymbol{\varphi}(1) = 0. \end{cases} \tag{2.8}$$

As a post-processing step, one can compute

$$\begin{cases} \boldsymbol{a}_- = \boldsymbol{\varphi}(0) - \boldsymbol{a}_+, \\ \boldsymbol{b}_+ = \boldsymbol{\varphi}(1)e^{-ik}, \end{cases} \tag{2.9}$$

which give the magnitudes of the waves traveling toward $-\infty$ and ∞, respectively.

To recap, we started with (2.1) which was applied when $x \in (0, L)$, while a simpler version was used in the rest of the space. Next, we wrote the constraints on the continuity and differentiability of the solution at the end points of $(0, L)$, after which scaling of x was performed in order to change the domain to $(0, 1)$. The equation was further simplified by renaming variables and defining new constants. Finally non-important constraint equations were eliminated and a concise set of equations (2.8) was obtained. The solution to (2.8) itself is not necessarily very interesting, but the relevant information can be computed via (2.9).

2.3 Electron Transport

Even though the scattered wave amplitudes (2.9) characterize the scattering process, a clearer way to present the same information are the transmission and reflection coefficients, also called amplitudes. For a single spin system, the transmission and reflection amplitudes would be

$$t = \frac{b_+}{a_+}, \quad r = \frac{a_-}{a_+},$$

respectively, and the corresponding probabilities for transmission and reflection

$$T = |t|^2, \quad R = |r|^2.$$

However, since we have two spin components, we are interested in, e.g., how down spin transmits as up spin. Such interactions happen due to the magnetic field. Therefore, the coefficients are represented as matrices

$$\boldsymbol{t} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}, \quad \boldsymbol{r} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix},$$

where the subscript 1 corresponds to up spin and 2 to down spin. The indices are arranged so that t_{12} represents the change of down spin to up spin. Analogously to the single spin case, the amplitudes $\boldsymbol{t}$ and $\boldsymbol{r}$ satisfy

$$\begin{cases} \boldsymbol{b}_+ = \boldsymbol{t}\boldsymbol{a}_+, \\ \boldsymbol{a}_- = \boldsymbol{r}\boldsymbol{a}_+. \end{cases} \tag{2.10}$$

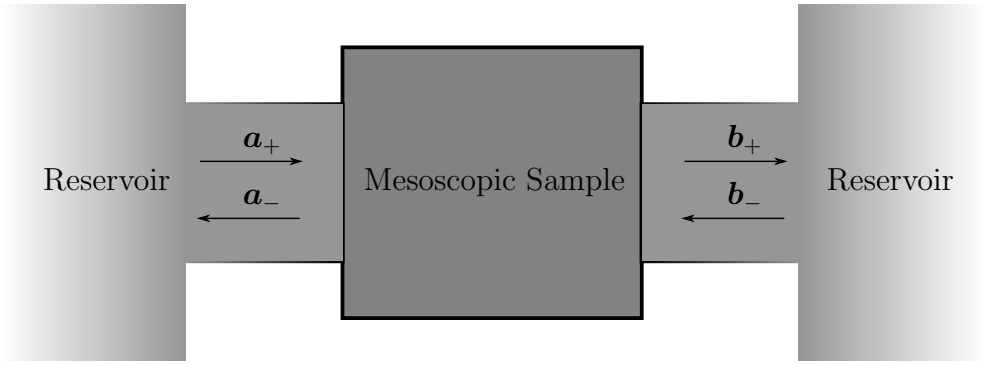

Figure 2.2: Mesoscopic sample connected to reservoirs by channels.

In the previous section we considered the abstract setup pictured in Figure 2.1. Here we examine a more practical configuration described in full detail in [2, ch. 7]; see Figure 2.2. The aim is to discuss the scattering matrix (S-matrix) and the concept of conductance in mesoscopic systems. This is intended only as an overview and we refer to [2] for the details. Roughly speaking, a mesoscopic sample is of small enough scale to exhibit quantum effects, but also large enough to enable treatment of some macroscopic concepts such as conductance. In a sense, a mesoscopic domain lies between microscopic and macroscopic. An example of mesoscopic systems would be a very thin metallic wire, or a semiconductor cooled to a temperature of order 1 Kelvin. There are many scales which need to be satisfied to enter the mesoscopic realm; they are listed in [2, ch. 7].

In the middle of the diagram in Figure 2.2, there is a mesoscopic sample. It is connected to macroscopic electron reservoirs via two leads which possibly contain multiple channels. The properties of the incoming electrons are determined by the physical conditions in the reservoirs, and additionally, the contact of the leads is expected not to cause reflections. This means that we may assume $\boldsymbol{a}_+$ and $\boldsymbol{b}_-$ known. In the S-matrix methodology we are not interested in wave functions in the mesoscopic sample, but instead the relation between input and output electrons is investigated. The transformation of the inputs, $\boldsymbol{a}_+$ and $\boldsymbol{b}_-$, into outputs, $\boldsymbol{a}_-$ and $\boldsymbol{b}_+$, is described by the S-matrix,

$$\boldsymbol{c}_{\text{out}} = \begin{bmatrix} \boldsymbol{a}_- \\ \boldsymbol{b}_+ \end{bmatrix} = \boldsymbol{S} \begin{bmatrix} \boldsymbol{a}_+ \\ \boldsymbol{b}_- \end{bmatrix} = \boldsymbol{S}\boldsymbol{c}_{\text{in}}. \tag{2.11}$$

It is easy to see that the first column of $\boldsymbol{S}$ can be written in the form

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{r} & ? \\ \boldsymbol{t} & ? \end{bmatrix}.$$

The second column is analogous since we could as well have chosen to fix $\boldsymbol{b}_-$ and set $\boldsymbol{a}_+ = 0$ when deriving (2.10). Let the corresponding transmission and reflection amplitudes be $\boldsymbol{t'}$ and $\boldsymbol{r'}$, respectively. The whole S-matrix then reads as

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{r} & \boldsymbol{t'} \\ \boldsymbol{t} & \boldsymbol{r'} \end{bmatrix}.$$

In order for the transmission and reflection probabilities to be conserved, the incoming flux $\|\boldsymbol{c}_{\text{in}}\|^2$ must equal the outgoing flux $\|\boldsymbol{c}_{\text{out}}\|^2$. In other words,

$$\boldsymbol{c}_{\text{in}}^* \boldsymbol{c}_{\text{in}} = \boldsymbol{c}_{\text{out}}^* \boldsymbol{c}_{\text{out}}.$$

Applying (2.11) to $\boldsymbol{c}_{\text{out}}$ gives

$$\boldsymbol{c}_{\text{in}}^* \boldsymbol{c}_{\text{in}} = \boldsymbol{c}_{\text{in}}^* \boldsymbol{S}^* \boldsymbol{S} \boldsymbol{c}_{\text{in}}$$

which simply means that $\boldsymbol{S}^* \boldsymbol{S} = \boldsymbol{I}$, i.e., $\boldsymbol{S}$ is unitary.

In the case of time reversal symmetry the S-matrix is also symmetric. Time reversal symmetry means that if $\boldsymbol{\Psi}(x,t)$ is a solution to the Schrödinger equation with some Hamiltonian H, then the time reversed and complex conjugated wave function $\overline{\boldsymbol{\Psi}}(x,-t)$ is also a solution for the same Hamiltonian. This happens when $\overline{H} = H$. To show this, consider (2.1) in the wave function formalism

$$H\boldsymbol{\Psi}(x,t) = i\hbar \frac{\partial}{\partial t} \boldsymbol{\Psi}(x,t). \tag{2.12}$$

Taking the complex conjugate gives

$$\overline{H\boldsymbol{\Psi}(x,t)} = \overline{H}\ \overline{\boldsymbol{\Psi}}(x,t) = -i\hbar \frac{\partial}{\partial t} \overline{\boldsymbol{\Psi}}(x,t).$$

Now changing the variable as $t \mapsto -t$ leads to

$$\overline{H}\ \overline{\boldsymbol{\Psi}}(x,-t) = i\hbar \frac{\partial}{\partial t} \overline{\boldsymbol{\Psi}}(x,-t)$$

if the Hamiltonian is not time-dependent. Hence $H = \overline{H}$, implies that

$$H\overline{\boldsymbol{\Psi}}(x,-t) = i\hbar \frac{\partial}{\partial t} \overline{\boldsymbol{\Psi}}(x,-t),$$

that is, $\overline{\boldsymbol{\Psi}}(x,-t)$ is also a solution to (2.12). When we have the time reversal property, implying the symmetry of $\boldsymbol{S}$, we know that $\boldsymbol{r} = \boldsymbol{r'}$ and $\boldsymbol{t} = \boldsymbol{t'}$, i.e., the transmission and reflection do not depend on the direction of the incoming wave.

Does our scattering problem have this symmetry? In quantum physics, the magnetic field is defined to be antisymmetric with respect to time reversal. This means that, if we reversed time, we would also have to change $\boldsymbol{B} \hookrightarrow -\boldsymbol{B}$. Since our model includes the magnetic field, it is not time reversal symmetric.

Scattering properties are directly related to observable electric conductance through the formula [6]

$$G = \frac{2e^2}{h} \operatorname{Tr}(\boldsymbol{t}^* \boldsymbol{t})$$

where e is the charge of an electron, h is the Planck constant and $\operatorname{Tr}(\cdot)$ is the trace of a matrix. This is the so-called Landauer–Büttiker formula. One can think that the conductance is the sum of the transmission possibilities of an electron. The quanta for the conductance is the coefficient

$$\frac{2e^2}{h}.$$

Moreover, the conductance is N times quantized if there are N channels with transmission probability one. The Landauer-Büttiker formula highlights the presence of quantum phenomena in mesoscopic systems.

Chapter 3

Direct Problem

In this chapter we formulate the scattering problem introduced in Chapter 2 in a more mathematical framework. The weak form of the boundary value problem (2.8) is derived in Section 3.1, and it is proven to have a unique solution in Section 3.2. Then, in Section 3.3 we build a numerical finite element solver for the problem. For more information on mathematics of scattering problems see, e.g., [12] and the references therein.

3.1 Weak Formulation

Let us start by writing the strong form of (2.8) in a slightly different manner:

$$\begin{cases} \boldsymbol{u}'' + \boldsymbol{F}\boldsymbol{u} = 0 \quad \text{in } (0,1), \\ \boldsymbol{u}'(0) + ik\boldsymbol{u}(0) = 2ik\boldsymbol{a}_+, \\ \boldsymbol{u}'(1) - ik\boldsymbol{u}(1) = 0. \end{cases} \tag{3.1}$$

We have renamed $\boldsymbol{\varphi}$ as $\boldsymbol{u}$ and also defined

$$\begin{aligned} \boldsymbol{F}(x) &= C\big((E - V(x)\boldsymbol{I} + \boldsymbol{\sigma} \cdot \boldsymbol{B}(x)\big) \\ &= \frac{2mL^2}{\hbar^2} \begin{bmatrix} E + B^3(x) - V(x) & B^1(x) - iB^2(x) \\ B^1(x) + iB^2(x) & E - B^3(x) - V(x) \end{bmatrix}. \end{aligned}$$

Recall that $k = \sqrt{CE}$ and $\boldsymbol{a}_+ \in \mathbb{C}$ is assumed to be known.

We obtain the weak formulation by multiplying both sides of (3.1) by a test function $\boldsymbol{v} \in [H^1(0,1)]^2$ and integrating over the domain

$$\int_0^1 \boldsymbol{u}'' \cdot \boldsymbol{v}\, dx + \int_0^1 (\boldsymbol{F}\boldsymbol{u}) \cdot \boldsymbol{v}\, dx = 0.$$

We leave the second integral as it is, but apply integration by parts to the first one, resulting in

$$\Big|_0^1 \boldsymbol{u}' \cdot \boldsymbol{v} - \int_0^1 \boldsymbol{u}' \cdot \boldsymbol{v}' \, dx + \int_0^1 (\boldsymbol{F}\boldsymbol{u}) \cdot \boldsymbol{v} \, dx = 0.$$

For the boundary terms we use the boundary conditions of (3.1) and get

$$-\int_0^1 \boldsymbol{u}' \cdot \boldsymbol{v}' \, dx + \int_0^1 (\boldsymbol{F}\boldsymbol{u}) \cdot \boldsymbol{v} \, dx + ik(\boldsymbol{u}(1) \cdot \boldsymbol{v}(1) + \boldsymbol{u}(0) \cdot \boldsymbol{v}(0)) = 2ik\boldsymbol{a}_+ \cdot \boldsymbol{v}(0),$$

after moving all $\boldsymbol{u}$-independent terms to the right hand side of the equation. Defining

$$(\boldsymbol{f}, \boldsymbol{g}) = \int_0^1 \boldsymbol{f} \cdot \boldsymbol{g} \, dx, \tag{3.2}$$

we can write the weak formulation in a compact form: Find $\boldsymbol{u} \in [H^1(0,1)]^2$ such that

$$\begin{aligned} -(\boldsymbol{u}', \boldsymbol{v}') + (\boldsymbol{F}\boldsymbol{u}, \boldsymbol{v}) + ik(\boldsymbol{u}(1) \cdot \boldsymbol{v}(1) + \boldsymbol{u}(0) \cdot \boldsymbol{v}(0)) &= 2ik\boldsymbol{a}_+ \cdot \boldsymbol{v}(0) \\ &\text{for all } \boldsymbol{v} \in [H^1(0,1)]^2. \end{aligned} \tag{3.3}$$

where $H^1(0,1)$ is a specific space of functions described in Section 3.2. We will discretize this weak form by FEM in Section 3.3, but first we need to check that the equation is solvable. In the next section we will show that (3.3) has a unique solution; it could actually be shown that (3.3) and (3.1) are equivalent if the derivatives in (3.1) are interpreted as weak ones [7].

3.2 Existence and Uniqueness of the Solution

The left hand side of (3.3) is known as the bilinear form $b(\boldsymbol{u}, \boldsymbol{v})$ and the right hand side as the load functional $l(\boldsymbol{v})$. We may thus rewrite (3.3) as

$$b(\boldsymbol{u}, \boldsymbol{v}) = l(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in [H^1(0,1)]^2 \tag{3.4}$$

for which we search a solution $\boldsymbol{u}$ in $[H^1(0,1)]^2$ which is a first order Sobolev space. A function belongs to this space if it is square integrable in the interval $(0,1)$ and its first weak derivative has also the same property. This allows us to consider terms like $(\boldsymbol{u}', \boldsymbol{v}')$. The Sobolev space in question is equipped with the norm

$$\|\boldsymbol{u}\|_{H^1(0,1)} = \|\boldsymbol{u}'\|_{L^2(0,1)} + \|\boldsymbol{u}\|_{L^2(0,1)} \tag{3.5}$$

where the L^2 norm is defined via the corresponding inner product:

$$(\boldsymbol{u}, \boldsymbol{v})_{L^2(0,1)} = \int_0^1 \boldsymbol{u}^* \boldsymbol{v} \, dx \tag{3.6}$$

and

$$\|\boldsymbol{u}\|_{L^2(0,1)} = \left(\int_0^1 |\boldsymbol{u}|^2 \, dx, \right)^{\frac{1}{2}}.$$

Finally, before moving on to the existence proof, notice that the problem (3.3) could as well be written with the help of the L^2 inner product instead of the pairing defined in (3.2), since we could equally well use the test function $\overline{\boldsymbol{v}}$ instead of $\boldsymbol{v}$. In the rest of this section, we assume that $b(\cdot, \cdot)$ is defined using $(\cdot, \cdot)_{L^2(0,1)}$ rather than $(\cdot, \cdot)$.

Normally, one would prove the existence and uniqueness of the solution to (3.4) by showing that

$$|b(\boldsymbol{u}, \boldsymbol{v})| \leq C \|\boldsymbol{u}\|_{H^1(0,1)} \|\boldsymbol{v}\|_{H^1(0,1)}, \tag{3.7}$$

$$|b(\boldsymbol{u}, \boldsymbol{u})| \geq C \|\boldsymbol{u}\|^2_{H^1(0,1)}, \tag{3.8}$$

for all $\boldsymbol{u}, \boldsymbol{v} \in [H^1(0,1)]^2$, and utilizing Lax–Milgram theorem [7]. In our case, (3.8) is not satisfied and thus a different approach is needed. For (3.8), we will actually prove a slightly modified version:

$$|b(\boldsymbol{u}, \boldsymbol{u})| + c\|\boldsymbol{u}\|^2_{L^2(0,1)} \geq C \|\boldsymbol{u}\|^2_{H^1(0,1)} \quad \text{for some } c > 0,$$

which turns out to be sufficient for the unique existence.

We start the proof by showing that there cannot be two or more solutions.

Lemma 3.1. *The problem* (3.1) *has at most one solution if the components of* $\boldsymbol{F}$ *are continuous and bounded, implying in particular that*

$$\|\boldsymbol{F}\boldsymbol{u}\|_{L^2(0,1)} \leq C \|\boldsymbol{u}\|_{L^2(0,1)},$$

and $\boldsymbol{F}$ *is Hermitian*

$$\boldsymbol{F}^* = \boldsymbol{F}.$$

Proof. Consider the original scattering problem (3.1) having two solutions, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, and denote $\boldsymbol{w} = \boldsymbol{u}_1 - \boldsymbol{u}_2$. Then the difference clearly satisfies

$$\begin{cases} \boldsymbol{w}'' + \boldsymbol{F}\boldsymbol{w} = 0 \quad \text{in } (0,1), \\ \boldsymbol{w}'(0) + ik\boldsymbol{w}(0) = 0, \\ \boldsymbol{w}'(1) - ik\boldsymbol{w}(1) = 0, \end{cases}$$

and obviously,

$$|\boldsymbol{w}'(0) + ik\boldsymbol{w}(0)|^2 + |\boldsymbol{w}'(1) - ik\boldsymbol{w}(1)|^2 = 0$$

This results in

$$\begin{aligned}\left(k^2|\boldsymbol{w}(0)|^2+|\boldsymbol{w}'(0)|^2\right) + \left(k^2|\boldsymbol{w}(1)|^2+|\boldsymbol{w}'(1)|^2\right)\\ +ik\left(\boldsymbol{w}'(0)^*\boldsymbol{w}(0)-\boldsymbol{w}(0)^*\boldsymbol{w}'(0)+\boldsymbol{w}(1)^*\boldsymbol{w}'(1)-\boldsymbol{w}'(1)^*\boldsymbol{w}(1)\right) = 0.\end{aligned} \tag{3.9}$$

From the identities

$$\begin{aligned}\boldsymbol{w}'(0)^*\boldsymbol{w}(0) - \boldsymbol{w}(0)^*\boldsymbol{w}'(0) &= 2i\,\mathrm{Im}\left(\boldsymbol{w}'(0)^*\boldsymbol{w}(0)\right),\\ \boldsymbol{w}(1)^*\boldsymbol{w}'(1) - \boldsymbol{w}'(1)^*\boldsymbol{w}(1) &= 2i\,\mathrm{Im}\left(\boldsymbol{w}'(1)^*\boldsymbol{w}(1)\right),\end{aligned}$$

it follows that in (3.9) we have the term

$$2i\,\mathrm{Im}\left(\boldsymbol{w}'(0)^*\boldsymbol{w}(0)\right) - 2i\,\mathrm{Im}\left(\boldsymbol{w}'(1)^*\boldsymbol{w}(1)\right) = -2i\,\mathrm{Im}\left(\Big|_0^1 \boldsymbol{w}'(x)^*\boldsymbol{w}(x)\right).$$

Thus (3.9) altogether becomes

$$\begin{aligned}\left(k^2|\boldsymbol{w}(0)|^2 + |\boldsymbol{w}'(0)|^2\right) + \left(k^2|\boldsymbol{w}(1)|^2 + |\boldsymbol{w}'(1)|^2\right)\\ + 2k\,\mathrm{Im}\left(\Big|_0^1 \boldsymbol{w}'(x)^*\boldsymbol{w}(x)\right) = 0.\end{aligned} \tag{3.10}$$

We proceed by showing that the last term in (3.10) is zero:

$$\begin{aligned}\Big|_0^1 \boldsymbol{w}'(x)^*\boldsymbol{w}(x) &= \int_0^1 \frac{d}{dx}\left(\boldsymbol{w}'(x)^*\boldsymbol{w}(x)\right)\,dx\\ &= \int_0^1 \left(\boldsymbol{w}''(x)^*\boldsymbol{w}(x) + \boldsymbol{w}'(x)^*\boldsymbol{w}'(x)\right)\,dx\\ &= \int_0^1 \left(\left(-\boldsymbol{F}\boldsymbol{w}(x)\right)^*\boldsymbol{w}(x) + |\boldsymbol{w}'(x)|^2\right)\,dx,\end{aligned}$$

where, in the last step, we have used the identity $\boldsymbol{w}'' + \boldsymbol{F}\boldsymbol{w} = 0$ (cf. (3.1)). Note that this integral makes sense since $\boldsymbol{F}$ is bounded. Because $|\boldsymbol{w}'(x)|^2$ is real, the imaginary part of the integral is simply

$$\mathrm{Im}\left(\Big|_0^1 \boldsymbol{w}'(x)^*\boldsymbol{w}(x)\right) = -\,\mathrm{Im}\left(\int_0^1 \left(\boldsymbol{F}\boldsymbol{w}(x)\right)^*\boldsymbol{w}(x)\,dx\right) \tag{3.11}$$

which is just the imaginary part of the L^2 inner product

$$-\operatorname{Im}\left((\boldsymbol{F}\boldsymbol{w},\boldsymbol{w})_{L^2(0,1)}\right).$$

However, since $\boldsymbol{F}$ is Hermitian,

$$(\boldsymbol{F}\boldsymbol{w},\boldsymbol{w})_{L^2(0,1)} = (\boldsymbol{w},\boldsymbol{F}\boldsymbol{w})_{L^2(0,1)} = \overline{(\boldsymbol{F}\boldsymbol{w},\boldsymbol{w})}_{L^2(0,1)}$$

which simply means that $(\boldsymbol{F}\boldsymbol{w},\boldsymbol{w})_{L^2(0,1)}$ is real. Consequently, due to (3.11),

$$\operatorname{Im}\left(\Big|_0^1 \boldsymbol{w}'(x)^*\boldsymbol{w}(x)\right) = 0$$

and thus (3.10) becomes

$$k^2|\boldsymbol{w}(0)|^2 + |\boldsymbol{w}'(0)|^2 + k^2|\boldsymbol{w}(1)|^2 + |\boldsymbol{w}'(1)|^2 = 0.$$

Because all of the terms above are non-negative they must actually be zero:

$$\boldsymbol{w}(0) = \boldsymbol{w}'(0) = \boldsymbol{w}(1) = \boldsymbol{w}'(1) = 0.$$

This means, in particular, that $\boldsymbol{w}$ satisfies the initial value problem

$$\begin{cases} \boldsymbol{w}''(x) = -\boldsymbol{F}\boldsymbol{w}(x), \\ \boldsymbol{w}(0) = 0, \\ \boldsymbol{w}'(0) = 0. \end{cases}$$

Due to Picard-Lindelöf theorem [4], it follows that $\boldsymbol{w} = 0$, i.e, $\boldsymbol{u}_1 = \boldsymbol{u}_2$. □

Theorem 3.2. *Under the assumptions of Lemma 3.1, the variational problem* (3.3) *has a unique solution in* $[H^1(0,1)]^2$.

Proof. Let us denote

$$b(\boldsymbol{u},\boldsymbol{v}) = -(\boldsymbol{u}',\boldsymbol{v}')_{L^2(0,1)} + (\boldsymbol{F}\boldsymbol{u},\boldsymbol{v})_{L^2(0,1)} + ik(\boldsymbol{u}(1)^*\boldsymbol{v}(1) + \boldsymbol{u}(0)^*\boldsymbol{v}(0)),$$
$$l(\boldsymbol{v}) = 2ik\boldsymbol{a}_+^*\boldsymbol{v}(0).$$

As discussed earlier, we can write the weak form (3.3) using L^2 inner products; as a consequence, we have changed the dot products to conjugate transposes.

Due to the trace theorem [7],

$$|l(\boldsymbol{v})| \le 2k|\boldsymbol{a}_+||\boldsymbol{v}(0)| \le 2k|\boldsymbol{a}_+|\|\boldsymbol{v}\|_{H^1(0,1)}. \tag{3.12}$$

This demonstrates that the linear load functional $l : [H^1(0,1)]^2 \to \mathbb{C}$ is bounded. Moreover, using the Cauchy–Schwarz inequality and the trace theorem, we obtain

$$\begin{aligned}|b(\boldsymbol{u},\boldsymbol{v})| \leq \|\boldsymbol{u}'\|_{L^2(0,1)}\|\boldsymbol{v}'\|_{L^2(0,1)} + \|\boldsymbol{F}\boldsymbol{u}\|_{L^2(0,1)}\|\boldsymbol{v}\|_{L^2(0,1)} \\ + C_1\|\boldsymbol{u}\|_{H^1(0,1)}\|\boldsymbol{v}\|_{H^1(0,1)},\end{aligned}$$

which can be further simplified by resorting to the boundedness of $\boldsymbol{F}$:

$$\begin{aligned}|b(\boldsymbol{u},\boldsymbol{v})| \leq \|\boldsymbol{u}'\|_{L^2(0,1)}\|\boldsymbol{v}'\|_{L^2(0,1)} + C_2\|\boldsymbol{u}\|_{L^2(0,1)}\|\boldsymbol{v}\|_{L^2(0,1)} \\ + C_1\|\boldsymbol{u}\|_{H^1(0,1)}\|\boldsymbol{v}\|_{H^1(0,1)}.\end{aligned} \tag{3.13}$$

From the definition of the $H^1(0,1)$ norm, it easily follows that (3.13) transforms into

$$|b(\boldsymbol{u},\boldsymbol{v})| \leq C\|\boldsymbol{u}\|_{H^1(0,1)}\|\boldsymbol{v}\|_{H^1(0,1)}. \tag{3.14}$$

Thus the sesquilinear form $b : [H^1(0,1)]^2 \times [H^1(0,1)]^2 \to \mathbb{C}$ is also bounded.

Next, we show that

$$|b(\boldsymbol{u},\boldsymbol{v})| + c\|\boldsymbol{u}\|^2_{L^2(0,1)} \geq C\|\boldsymbol{u}\|^2_{H^1(0,1)}. \tag{3.15}$$

Naturally,

$$\begin{aligned}|b(\boldsymbol{u},\boldsymbol{u})| &\geq -\operatorname{Re}\left(b(\boldsymbol{u},\boldsymbol{u})\right) \\ &= \|\boldsymbol{u}'\|^2_{L^2(0,1)} - \operatorname{Re}\left((\boldsymbol{F}\boldsymbol{u},\boldsymbol{u})_{L^2(0,1)}\right) \\ &\geq \|\boldsymbol{u}'\|^2_{L^2(0,1)} - \left|(\boldsymbol{F}\boldsymbol{u},\boldsymbol{u})_{L^2(0,1)}\right|.\end{aligned}$$

Because by Cauchy–Schwarz and boundedness of $\boldsymbol{F}$

$$-\left|(\boldsymbol{F}\boldsymbol{u},\boldsymbol{u})_{L^2(0,1)}\right| \geq -C_1\|\boldsymbol{u}\|^2_{L^2(0,1)},$$

we may further estimate that

$$\begin{aligned}|b(\boldsymbol{u},\boldsymbol{u})| &\geq \|\boldsymbol{u}'\|^2_{L^2(0,1)} - C_1\|\boldsymbol{u}\|^2_{L^2(0,1)} \\ &= \|\boldsymbol{u}'\|^2_{L^2(0,1)} + \|\boldsymbol{u}\|^2_{L^2(0,1)} - (C_1+1)\|\boldsymbol{u}\|^2_{L^2(0,1)}.\end{aligned} \tag{3.16}$$

As a final step, we compute

$$\|\boldsymbol{u}'\|^2_{L^2(0,1)} + \|\boldsymbol{u}\|^2_{L^2(0,1)} \geq \frac{1}{2}\left(\|\boldsymbol{u}'\|_{L^2(0,1)} + \|\boldsymbol{u}\|_{L^2(0,1)}\right)^2 = C\|\boldsymbol{u}\|^2_{H^1(0,1)}. \tag{3.17}$$

Now, combining (3.16) and (3.17), we finally have

$$|b(\boldsymbol{u},\boldsymbol{u})| \geq C\|\boldsymbol{u}\|^2_{H^1(0,1)} - c\|\boldsymbol{u}\|^2_{L^2(0,1)}$$

and subsequently the modified coercivity result (3.15) holds.

From (3.12), (3.14) and (3.15), it can be deduced by the Lax–Milgram lemma and Fredholm theory [3, ch. 5] that (3.3) is uniquely solvable if and only if it admits at most one solution. Thus the claim follows from Lemma 3.1. □

Remark that in our case the components of $\boldsymbol{F}$ are clearly continuous and bounded because any physical potential or magnetic field satisfies such conditions.

3.3 Finite Element Solver

In this section we will introduce a FEM discretization for (3.3) and provide notes on the numerical implementation. Observe that in addition to the solution of (3.3), we also need to compute the transmission and reflection amplitudes (2.10). Moreover, the amplitudes themselves are not the most important aspects regarding the aim of the thesis: The actual transmission and reflection probabilities are computed using the amplitudes.

We start by discretizing (3.3). Let x be divided in $N-1$ elements, that is, we have N nodes for x. We employ a uniform spacing of the grid points. Our mesh parameter is thus defined as

$$h = \frac{1}{N-1}.$$

As the basis functions we use $\boldsymbol{v}_k \in V_h$ defined via

$$\begin{cases} \boldsymbol{v}_k &= \begin{bmatrix} \phi_k \\ 0 \end{bmatrix}, \quad 1 \leq k \leq N, \\ \boldsymbol{v}_k &= \begin{bmatrix} 0 \\ \phi_{k-N} \end{bmatrix}, \quad N+1 \leq k \leq 2N, \end{cases}$$

where ϕ_k is the standard one-dimensional 'tent function'

$$\phi_k(x) = \begin{cases} \dfrac{x - x_{k-1}}{x_k - x_{k-1}}, & x \in [x_{k-1}, x_k], \\ \dfrac{x_{k+1} - x}{x_{k+1} - x_k}, & x \in [x_k, x_{k+1}], \\ 0, & \text{otherwise}, \end{cases}$$

with the obvious modification for ϕ_1 and ϕ_N. Since we have $2N$ basis functions our FEM system will be of size $2N \times 2N$. However, since the one-dimensional basis functions ϕ_i are the same for both up and down spins, the system to be solved can be written as a block matrix

$$\boldsymbol{A}\boldsymbol{u}_h = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \end{bmatrix} = \boldsymbol{f}, \tag{3.18}$$

where $\boldsymbol{A}$ corresponds to the left hand side of (3.3) and $\boldsymbol{f}$ to the right hand side. The solution of (3.18) is denoted as $\boldsymbol{u}_h = \left[\boldsymbol{u}_1^T, \boldsymbol{u}_2^T\right]^T$ where $\boldsymbol{u}_1$ corresponds to up spin values and $\boldsymbol{u}_2$ to down spin. In other words, the vector $\boldsymbol{u}_h$ first contains discretized up spin component of $\boldsymbol{u}$ and then the down spin component. The system matrix $\boldsymbol{A}$ can be written as a sum of three parts:

1. $\boldsymbol{K}$ corresponding to $-(\boldsymbol{u}', \boldsymbol{v}')$,
2. $\boldsymbol{V}$ corresponding to $(\boldsymbol{F}\boldsymbol{u}, \boldsymbol{v})$,
3. the boundary conditions.

These will be divided into similar block matrices as in (3.18). Note that the off-diagonal blocks describe how spins change. For example, $\boldsymbol{A}_{12}$ describes how down spin transforms into up spin in the scattering process.

We start by introducing $\boldsymbol{K}$. Writing

$$\boldsymbol{u} = \sum_{k=1}^{2N} u_k \boldsymbol{v}_k$$

and considering only one block of $\boldsymbol{K}$ at a time, enables us to work with scalar functions instead of vector valued ones. Only the scalar representation of $\boldsymbol{u}$ needs to be considered:

$$u = \sum_{j=1}^{N} u_j \phi_j.$$

Note that depending on the block we want to compute, u_j is indexed from 1 to N or from $N+1$ to $2N$. The one-dimensional basis functions will always have indices running from 1 to N. With the new representation, the first integral of (3.3) takes the form

$$-(\boldsymbol{u}', \boldsymbol{v}') = -\int_0^1 \sum_{j=1}^{N} u_j \phi_j' \phi_i' \, dx = \sum_{j=1}^{N} \left(-\int_0^1 \phi_i' \phi_j' \, dx \right) u_j.$$

This is essentially a matrix vector product $\tilde{\boldsymbol{K}}\tilde{\boldsymbol{u}}$, where

$$\tilde{K}_{ij} = -\int_0^1 \phi_i' \phi_j' \, dx, \quad i, j = 1, \ldots, N,$$

and 'tildes' indicate that the matrix and the vector are blocks of $\boldsymbol{K}$ and $\boldsymbol{u}_h$, respectively. To determine the elements of $\tilde{\boldsymbol{K}}$, we consider a subintegral

$$-\int_{x_k}^{x_{k+1}} \phi_i' \phi_j' \, dx$$

which takes the values

$$\begin{cases} -h^{-1}, & i = j = k, \\ h^{-1}, & i = k,\ j = k+1, \\ h^{-1}, & i = k+1,\ j = k, \\ -h^{-1}, & i = j = k+1, \\ 0, & \text{otherwise}, \end{cases}$$

where h is the mesh parameter. Summing over k, we get the whole integral

$$\tilde{K}_{ij} = -\int_0^1 \phi_i' \phi_j' \, dx = \begin{cases} -\dfrac{1}{h}, & i = j = 1, \\ -\dfrac{1}{h}, & i = j = N, \\ -\dfrac{2}{h}, & i = j \neq 1 \text{ or } N, \\ \dfrac{1}{h}, & j = i + 1, \\ \dfrac{1}{h}, & j = i - 1, \\ 0, & \text{otherwise}. \end{cases}$$

In matrix form this is

$$\tilde{\boldsymbol{K}} = \frac{1}{h} \begin{bmatrix} -1 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -1 \end{bmatrix}.$$

Since $-(\boldsymbol{u}', \boldsymbol{v}')$ does not have any spin interaction terms, $\boldsymbol{K}$ is block diagonal:

$$\boldsymbol{K} = \begin{bmatrix} \tilde{\boldsymbol{K}} & \\ & \tilde{\boldsymbol{K}} \end{bmatrix}. \tag{3.19}$$

This is the first part of our FEM system matrix.

Next, we calculate $\boldsymbol{V}$ which, like $\boldsymbol{K}$, consists of four blocks. The blocks are built similarly to $\tilde{\boldsymbol{K}}$, but this time the general form of the integrals is slightly different:

$$\int_0^1 g(x)\phi_i(x)\phi_j(x)\, dx. \tag{3.20}$$

This is due to the weak formulation (3.3) containing the inner product of $\boldsymbol{Fu}$ and $\boldsymbol{v}$. To approximate the integral (3.20), we represent g in the discretized basis,

$$g(x) = \sum_{k=1}^{N} g_k \phi_k.$$

Restricting our attention to the k^{th} interval gives

$$\int_{x_k}^{x_{k+1}} g(x)\phi_i\phi_j\, dx = \int_{x_k}^{x_{k+1}} (g_k\phi_k + g_{k+1}\phi_{k+1})\, \phi_i\phi_j\, dx$$

because ϕ_l is zero when l is not k or $k+1$. The same is also required of i and j. Remembering that the length of the interval $[x_k, x_{k+1}]$ is h, we can evaluate the above integral explicitly,

$$\int_{x_k}^{x_{k+1}} (g_k\phi_k + g_{k+1}\phi_{k+1})\, \phi_i\phi_j\, dx = \begin{cases} g_k\dfrac{h}{4} + g_{k+1}\dfrac{h}{12}, & i = j = k, \\ g_k\dfrac{h}{12} + g_{k+1}\dfrac{h}{12}, & i = k,\ j = k+1, \\ g_k\dfrac{h}{12} + g_{k+1}\dfrac{h}{12}, & i = k+1,\ j = k, \\ g_k\dfrac{h}{12} + g_{k+1}\dfrac{h}{4}, & i = j = k+1. \end{cases}$$

Summing over k from 1 to $N-1$ gives us the whole integral from 0 to 1,

$$\int_0^1 g(x)\phi_i\phi_j\,dx = \begin{cases} \frac{h}{12}\left(3g_1+g_2\right), & i=j=1, \\ \frac{h}{12}\left(g_{N-1}+3g_N\right), & i=j=N, \\ \frac{h}{12}\left(g_{i-1}+6g_i+g_{i+1}\right), & i=j\neq 1 \text{ or } N, \\ \frac{h}{12}\left(g_i+g_{i+1}\right), & j=i+1, \\ \frac{h}{12}\left(g_{i-1}+g_i\right), & j=i-1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.21}$$

We will next apply this formula to $(\boldsymbol{F}\boldsymbol{u},\boldsymbol{v})$.

As earlier, in a single block matrix, the test function $\boldsymbol{v}$ is replaced by ϕ_i, and we denote the relevant components of $\boldsymbol{u}$ and $\boldsymbol{F}$ by

$$u=\sum_{j=1}^{N} u_j\phi_j$$

and $f(x)$, respectively. Then,

$$(\boldsymbol{F}\boldsymbol{u},\boldsymbol{v}) = \int_0^1 f(x)\sum_{j=1}^{N} u_j\phi_j\phi_i\,dx = \sum_{j=1}^{N}\int_0^1 f(x)\phi_i\phi_j\,dx\;u_j$$

which can be written as a matrix vector product

$$\tilde{\boldsymbol{V}}\tilde{\boldsymbol{u}} = \sum_{j=1}^{N}\tilde{V}_{ij}u_j$$

where $\tilde{\boldsymbol{V}}$ is a block of $\boldsymbol{V}$ defined componentwise via

$$\tilde{V}_{ij} = \int_0^1 f(x)\phi_i\phi_j\,dx.$$

Recalling (3.1), it is easy to conclude the blockwise definitions

$$f(x) = \begin{cases} E+B^3(x)-V(x), & \text{for } \boldsymbol{V}_{11}, \\ B^1(x)-iB^2(x), & \text{for } \boldsymbol{V}_{12}, \\ B^1(x)+iB^2(x), & \text{for } \boldsymbol{V}_{21}, \\ E-B^3(x)-V(x), & \text{for } \boldsymbol{V}_{22}, \end{cases} \tag{3.22}$$

and altogether that

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{V}_{11} & \boldsymbol{V}_{12} \\ \boldsymbol{V}_{21} & \boldsymbol{V}_{22} \end{bmatrix} \tag{3.23}$$

which is formed by plugging the functions (3.22) as g in (3.21).

Finally, we include the boundary conditions in the discretized systems. Those are represented by the term

$$ik(\boldsymbol{u}(1) \cdot \boldsymbol{v}(1) + \boldsymbol{u}(0) \cdot \boldsymbol{v}(0)) \tag{3.24}$$

in (3.3). We incorporate the term (3.24) in $\boldsymbol{V}$ of (3.23). The needed changes are simply

$$\begin{array}{rcl} V_{11}^{11} & = & V_{11}^{11} + ik, \\ V_{11}^{NN} & = & V_{11}^{NN} + ik, \\ V_{22}^{11} & = & V_{22}^{11} + ik, \\ V_{22}^{NN} & = & V_{22}^{NN} + ik, \end{array} \tag{3.25}$$

where the subscripts indicate the matrix block and the superscripts the element of the block in question.

The system matrix $\boldsymbol{A}$ is finally obtained by combining (3.19), (3.23) and (3.25):

$$\boldsymbol{A} = \boldsymbol{K} + \boldsymbol{V}.$$

Solving the approximative wave function is now straightforward:

$$\boldsymbol{u}_h = \boldsymbol{A}^{-1}\boldsymbol{f}$$

where $\boldsymbol{f}$ of (3.18) is created from the right hand side of (3.3) as

$$\boldsymbol{f} = \begin{bmatrix} 2ika_+^1 \\ 0 \\ \vdots \\ 0 \\ 2ika_+^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \boldsymbol{f}_1 \\ \\ \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \boldsymbol{f}_2 \end{matrix} . \tag{3.26}$$

The existence of the inverse matrix $\boldsymbol{A}^{-1}$ is a subtle issue; see [10] for more details and related stability estimates. Here, a_+^1 and a_+^2 are the components of the incoming wave amplitude

$$\boldsymbol{a}_+ = \begin{bmatrix} a_+^1 \\ a_+^2 \end{bmatrix}.$$

When deducing the transmission and reflection amplitudes, some difficulties occur because both $\boldsymbol{t}$ and $\boldsymbol{r}$ have four elements. For example solving all elements of $\boldsymbol{t}$ from

$$\boldsymbol{b}_+ = \boldsymbol{t}\boldsymbol{a}_+$$

cannot be performed directly. Fortunately, the components of $\boldsymbol{t}$ and $\boldsymbol{r}$ can be found by dividing the challenge in two parts. Since the weak formulation is linear with respect to $\boldsymbol{a}_+$, we can solve each column of $\boldsymbol{t}$ and $\boldsymbol{r}$ by running our solver in two batches. In the first batch we choose $\boldsymbol{a}_+$ to be

$$\boldsymbol{a}_+ = \begin{bmatrix} a_+^1 \\ 0 \end{bmatrix} \tag{3.27}$$

and in the second

$$\boldsymbol{a}_+ = \begin{bmatrix} 0 \\ a_+^2 \end{bmatrix}. \tag{3.28}$$

Applying $\boldsymbol{t}$ to the first $\boldsymbol{a}_+$ gives

$$\boldsymbol{t}\begin{bmatrix} a_+^1 \\ 0 \end{bmatrix} = \begin{bmatrix} t_{11}a_+^1 \\ t_{21}a_+^1 \end{bmatrix} = \boldsymbol{b}_+$$

and thus

$$t_{11} = \frac{b_+^1}{a_+^1}, \quad t_{21} = \frac{b_+^2}{a_+^1}. \tag{3.29}$$

Changing b_+ to a_- gives the first column of $\boldsymbol{r}$ since $\boldsymbol{a}_- = \boldsymbol{r}\boldsymbol{a}_+$ by (2.10). When running the second batch, the other column is obtained as

$$t_{12} = \frac{b_+^1}{a_+^2}, \quad t_{22} = \frac{b_+^2}{a_+^2}. \tag{3.30}$$

Again, making the appropriate change of b_+ to a_- gives the last column of $\boldsymbol{r}$. Recall that the transmitted and reflected wave amplitudes $\boldsymbol{b}_+$ and $\boldsymbol{a}_-$ are computed using (2.9), re-written here in the notation of this chapter:

$$\begin{cases} \boldsymbol{a}_- = \boldsymbol{u}(0) - \boldsymbol{a}_+, \\ \boldsymbol{b}_+ = \boldsymbol{u}(1)e^{-ik}. \end{cases} \tag{3.31}$$

We can now sketch out our numerical solver. Ultimately, we are not interested in a single set of transmission amplitudes but instead in their relation to the energy of the incoming wave, E. Furthermore, the amplitudes themselves are not the values we will be working with: We will consider real-valued transmission and reflection probabilities

$$\begin{cases} T_{ij} = |t_{ij}|^2, \\ R_{ij} = |r_{ij}|^2, \end{cases} \tag{3.32}$$

as the data for the inverse scattering problem in Chapter 4. This, however, loses the phase information on the complex amplitudes t_{ij} and r_{ij}. Therefore we describe here a solver for the amplitudes as it is more general; the change to real probabilities is just a straightforward computation.

We have to repeat the solving process multiple times as the FEM system provides the solution only for a single energy at a time. Let the set of energies for which we want to solve the transmission coefficients be $\mathcal{E} \subset \mathbb{R}_+$, and let $\mathcal{A}$ denote the set of the two $\boldsymbol{a}_+$, (3.27) and (3.28), used in solving t_{ij} and r_{ij}, namely,

$$\mathcal{A} = \left\{ \begin{bmatrix} a_+^1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ a_+^2 \end{bmatrix} \right\}.$$

The program for computing $\boldsymbol{u}(x)$, $\boldsymbol{r}(E)$ and $\boldsymbol{t}(E)$ is presented in Algorithm 3.1. Some observations on the program are in order:

- $\boldsymbol{K}$, $\boldsymbol{V}_{12}$ and $\boldsymbol{V}_{21}$ can be created before looping over the energy levels because they do not depend on E.
- $\boldsymbol{V}_{11}$ and $\boldsymbol{V}_{22}$ depend on the energy via (3.22) and thus are updated during each energy iteration.
- The problem is linear with respect to $\boldsymbol{a}_+$. Therefore, we can compute the solution $\boldsymbol{u}_h$ as a linear combination of two components, each corresponding to a single $\boldsymbol{a}_+ \in \mathcal{A}$.

3.4 Sensitivity of the Equation

In this section we examine how the solution of (3.1) changes when $\boldsymbol{F}(x)$ is altered. Furthermore, we are interested in the changes in the transmission and reflection coefficients $\boldsymbol{t}$ and $\boldsymbol{r}$. The main idea is to add a small perturbation to $\boldsymbol{F}$ in the weak formulation (3.3). Resorting to an approximative version of this new equation, one can determine the sensitivity, or the functional derivate, with respect to $\boldsymbol{F}$.

The sensitivity is important when solving the inverse scattering problem, i.e., finding some components of $\boldsymbol{F}$ from, say, the knowledge of transmission coefficients. This need originates from the dependence of $\boldsymbol{u}$ on $\boldsymbol{F}$ being nonlinear in (3.1): the required components of $\boldsymbol{F}$ cannot be solved directly but instead an iterative scheme is needed.

Let us give an outline of the derivation of the weak formulation for solving the functional derivative. We start by considering (3.3) and changing $\boldsymbol{F}$ to $\boldsymbol{F} + \varepsilon(x)\boldsymbol{I}$, that is, we are only interested in changes of the diagonal. This choice is made since the quantity of interest in our inverse problem

(cf. Chapter 4) is the potential V which only affects the diagonal of $\boldsymbol{F}$. The sensitivity with respect to the magnetic field $\boldsymbol{B}$ is not considered here. Note that the perturbation parameter is a function on x, representing a small change in the potential function $V(x)$.

Let $\tilde{\boldsymbol{u}}$ be the solution to the perturbed weak formulation, that is,

$$-(\tilde{\boldsymbol{u}}', \boldsymbol{v}') + ((\boldsymbol{F} + \varepsilon \boldsymbol{I})\tilde{\boldsymbol{u}}, \boldsymbol{v}) + ik(\tilde{\boldsymbol{u}}(1) \cdot \boldsymbol{v}(1) + \tilde{\boldsymbol{u}}(0) \cdot \boldsymbol{v}(0)) = 2ik\boldsymbol{a}_+ \cdot \boldsymbol{v}(0)$$

which can be written as

$$\begin{aligned} -(\tilde{\boldsymbol{u}}', \boldsymbol{v}') + (\boldsymbol{F}\tilde{\boldsymbol{u}}, \boldsymbol{v}) + (\varepsilon\tilde{\boldsymbol{u}}, \boldsymbol{v}) + ik(\tilde{\boldsymbol{u}}(1) \cdot \boldsymbol{v}(1) + \tilde{\boldsymbol{u}}(0) \cdot \boldsymbol{v}(0)) \\ = 2ik\boldsymbol{a}_+ \cdot \boldsymbol{v}(0). \end{aligned} \tag{3.33}$$

Next, we subtract the solution of the original weak formulation (3.3) from (3.33) and obtain

$$\begin{aligned} -(\tilde{\boldsymbol{u}}' - \boldsymbol{u}', \boldsymbol{v}') + (\boldsymbol{F}(\tilde{\boldsymbol{u}} - \boldsymbol{u}), \boldsymbol{v}) + ik((\tilde{\boldsymbol{u}}(1) - \boldsymbol{u}(1)) \cdot \boldsymbol{v}(1) + (\tilde{\boldsymbol{u}}(0) - \boldsymbol{u}(0)) \cdot \boldsymbol{v}(0)) \\ = -(\varepsilon\tilde{\boldsymbol{u}}, \boldsymbol{v}). \end{aligned}$$

Setting

$$\boldsymbol{w} = \tilde{\boldsymbol{u}} - \boldsymbol{u}, \tag{3.34}$$

the above formula simplifies to

$$-(\boldsymbol{w}', \boldsymbol{v}') + (\boldsymbol{F}\boldsymbol{w}, \boldsymbol{v}) + ik(\boldsymbol{w}(1) \cdot \boldsymbol{v}(1) + \boldsymbol{w}(0) \cdot \boldsymbol{v}(0)) = -(\varepsilon\tilde{\boldsymbol{u}}, \boldsymbol{v}). \tag{3.35}$$

Comparing this to (3.3), we can immediately see that only the right hand side of the equation has changed. There are two things worth noticing:

- Solving (3.35) does not make sense because it requires the knowledge of $\tilde{\boldsymbol{u}}$, and thus one could obtain $\boldsymbol{w}$ trivially by using (3.34).
- The function $\boldsymbol{w}$ cannot be a derivative of $\boldsymbol{u}$ with respect to ε since its dependence on ε is non-linear.

To fix this, we approximate $\tilde{\boldsymbol{u}}$.

One can show that the solution to the perturbed problem satisfies

$$\tilde{\boldsymbol{u}} = \boldsymbol{u} + O(\varepsilon),$$

in the appropriate function space topology, cf. [8]. With this in mind, we approximate

$$\begin{aligned} -(\varepsilon\tilde{\boldsymbol{u}}, \boldsymbol{v}) &= -(\varepsilon(\boldsymbol{u} + O(\varepsilon)), \boldsymbol{v}) \\ &= -(\varepsilon\boldsymbol{u} + O(\varepsilon^2), \boldsymbol{v}) \\ &= -(\varepsilon\boldsymbol{u}, \boldsymbol{v}) + O(\varepsilon^2). \end{aligned} \tag{3.36}$$

Motivated by (3.36), we define $D\boldsymbol{u}_\varepsilon$ to be the solution of a slightly altered version of (3.35):

$$\begin{aligned} -((D\boldsymbol{u}_\varepsilon)', \boldsymbol{v}') + (\boldsymbol{F}(D\boldsymbol{u}_\varepsilon), \boldsymbol{v}) + ik(D\boldsymbol{u}_\varepsilon(1) \cdot \boldsymbol{v}(1) + D\boldsymbol{u}_\varepsilon(0) \cdot \boldsymbol{v}(0)) \\ = -(\varepsilon \boldsymbol{u}, \boldsymbol{v}). \end{aligned} \tag{3.37}$$

The unique solvability of (3.37) in $[H^1(0,1)]^2$ follows exactly as that of (3.3) in Theorem 3.2 assuming that ε is bounded. Notice that $D\boldsymbol{u}_\varepsilon$ depends linearly on ε. Denoting $\tilde{\boldsymbol{w}} = \boldsymbol{w} - D\boldsymbol{u}_\varepsilon$ and utilizing (3.35), (3.36) and (3.37), we have

$$-(\tilde{\boldsymbol{w}}', \boldsymbol{v}') + (\boldsymbol{F}\tilde{\boldsymbol{w}}, \boldsymbol{v}) + ik(\tilde{\boldsymbol{w}}(1) \cdot \boldsymbol{v}(1) + \tilde{\boldsymbol{w}}(0) \cdot \boldsymbol{v}(0)) = O(\varepsilon^2). \tag{3.38}$$

Relying on properties of the bilinear form defined by left hand side, it could be shown that $\tilde{\boldsymbol{w}}$ behaves as a function of ε similarly to the right hand side of (3.38), that is,

$$\tilde{\boldsymbol{w}} = O(\varepsilon^2).$$

This clearly implies that $\varepsilon \mapsto D\boldsymbol{u}_\varepsilon$ is the best linear approximation of $\varepsilon \mapsto \boldsymbol{w}$. In other words, $D\boldsymbol{u}_\varepsilon$ is the Gateaux derivative [1, appx. A] of $\boldsymbol{u}$ in the direction ε.

When discretized, $D\boldsymbol{u}$ will be the approximate Jacobian with respect to elements of discretized diagonal components of $\boldsymbol{F}$, namely the potential $V(x)$. One can interpret this statement as

$$\boldsymbol{u}(\boldsymbol{F} + \Delta\boldsymbol{F}) \approx \boldsymbol{u}(\boldsymbol{F}) + D\boldsymbol{u}_{\Delta\boldsymbol{F}}(\boldsymbol{F})$$

if one regards the solution $\boldsymbol{u}$ to be a function of $\boldsymbol{F}$. The above derivation lacks mathematical rigor, but instead of being a proper mathematical proof, it is intended only as a motivation for the algorithm introduced in the next section.

3.5 Functional Derivative Solver

We start by discretizing the perturbation $\varepsilon(x)$. The same basis as earlier is used,

$$\varepsilon(x) = \sum_{i=1}^{N} \varepsilon_i \phi_i, \quad \varepsilon_i \geq 0, \quad i = 1, \ldots, N.$$

We choose to perturb a single component of the discretized V at a time, meaning that

$$\sum_{i=1}^{N} \varepsilon_i \phi_i = \varepsilon_l \phi_l.$$

Inserting this in (3.37) results in

$$-((D\boldsymbol{u}_\varepsilon)', \boldsymbol{v}') + (\boldsymbol{F}(D\boldsymbol{u}_\varepsilon), \boldsymbol{v}) + ik(D\boldsymbol{u}_\varepsilon(1)\cdot\boldsymbol{v}(1) + D\boldsymbol{u}_\varepsilon(1)\cdot\boldsymbol{v}(0)) = -\varepsilon_l(\phi_l\boldsymbol{u}, \boldsymbol{v}).$$

Dividing by ε_l gives us the derivative of $\boldsymbol{u}$ with respect to the l^{th} component of V:

$$-((D_l\boldsymbol{u})', \boldsymbol{v}') + (\boldsymbol{F}(D_l\boldsymbol{u}), \boldsymbol{v}) + ik(D_l\boldsymbol{u}(1)\cdot\boldsymbol{v}(1) + D_l\boldsymbol{u}(0)\cdot\boldsymbol{v}(0)) = -(\phi_l\boldsymbol{u}, \boldsymbol{v}).$$

The variational derivative problem is identical to the original direct FEM problem (3.3) except for the right hand side which has changed from $2ik\boldsymbol{a}_+ \cdot \boldsymbol{v}(0)$ to $-(\phi_l\boldsymbol{u}, \boldsymbol{v})$. This means that the derivative $D_l\boldsymbol{u}$ can be computed using the same program if the load $\boldsymbol{f}$ is modified. Of course, when computing all the derivatives, we must loop over l and store the results. We will first present the needed changes and then the resulting Algorithm 3.2.

Computing the load for the derivative problem is straightforward since

$$-(\phi_l\boldsymbol{u}, \boldsymbol{v}_j) = -\int_0^1 \phi_l\boldsymbol{u}\cdot\boldsymbol{v}_j\,dx$$

which is divided into two parts corresponding to the two spins, and the exact solution of (3.1) is replaced by the associated FEM approximation $\boldsymbol{u}_h$. This leads to integrals of the form

$$-\int_0^1 u(x)\phi_l\phi_j\,dx \tag{3.39}$$

which are simply evaluated using formula (3.21). It is important to notice that the formulation of the derivative problem requires the solution to the original problem. Furthermore, the division in two cases introduces the same blocks as for the direct FEM program. Thus, there will be two loads $\boldsymbol{y}_l^1$ and $\boldsymbol{y}_l^2$ in the discrete system

$$\boldsymbol{A}(D_l\boldsymbol{u}_h) = \begin{bmatrix}\boldsymbol{y}_l^1\\ \boldsymbol{y}_l^2\end{bmatrix} = \boldsymbol{y}_l \tag{3.40}$$

for the derivative.

The derivatives of $\boldsymbol{t}$ and $\boldsymbol{r}$ need also to be computed. A similar procedure as for the original problem is used. Looking at (3.31), we deduce that

$$\begin{cases} D\boldsymbol{a}_- = D\boldsymbol{u}(0),\\ D\boldsymbol{b}_+ = D\boldsymbol{u}(1)e^{-ik}. \end{cases} \tag{3.41}$$

In addition, (3.29) and (3.30) are changed appropriately to account for $D\boldsymbol{a}_-$ and $D\boldsymbol{b}_+$ instead of $\boldsymbol{a}_-$ and $\boldsymbol{b}_+$. The change is trivial since the equations are linear in $\boldsymbol{a}_-$ and $\boldsymbol{b}_+$. E.g.,

$$\begin{cases} Dt_{11} = \dfrac{Db_+^1}{a_+^1}, \\ Dt_{21} = \dfrac{Db_+^2}{a_+^1}. \end{cases} \tag{3.42}$$

We can now write the program for computing the derivatives with respect to V. It can be found in Algorithm 3.2. Notice that the derivatives

- are computed for each energy level E,
- have a component for each perturbed point in the discretized $V(x)$.

The derivatives inherit the form of their non-differentiated counterparts. For example, $D\boldsymbol{u}_h$ is discretized in space. This means that for every grid point, say x_i, there is a matrix

$$\boldsymbol{J} = [D\boldsymbol{u}_h]_i$$

which has components

$$J_{jl} = \frac{\partial u_{ij}}{\partial V_l}$$

where u_{ij} is $\boldsymbol{u}_h$ evaluated at x_i (for either spin) with the energy being at index j, and V_l denotes the perturbation of V at x_l. There is such a matrix for each spin component. The matrix $\boldsymbol{J} = \boldsymbol{J}(i)$ is the Jacobian of $\boldsymbol{u}_h$ at x_i over the energies E_j with respect to perturbation in V.

To get more insight, consider for instance the component t_{11} of $\boldsymbol{t}$. The Jacobian of t_{11} is given componentwise as

$$J_{jl} = \frac{\partial t_{11}(E_j)}{\partial V_l}$$

if one considers t_{11} as a function of the energy. In this sense

$$Dt_{11} = \frac{\partial t_{11}}{\partial V}.$$

Changing the discretized V to the perturbed version $V + \Delta V$ yields a new transmission amplitude with

$$t_{11,j}^{\text{new}} \approx t_{11,j} + \sum_{l=1}^{N} J_{jl} \, [\Delta V]_l$$

where $t_{11,j} = t_{11}(E_j)$. Written in vector form this reads

$$\boldsymbol{t}_{11}^{\text{new}} \approx \boldsymbol{t}_{11} + \boldsymbol{J}\Delta\boldsymbol{V}.$$

Derivatives of the transmission and reflection probabilities (3.32) can be computed with the chain rule. As previously, when considering the transmission coefficient

$$t_{ij,k} = t_{ij}(E_k)$$

and its derivative

$$Dt_{ij,kl} = \frac{\partial t_{ij}(E_j)}{\partial V_l},$$

we may write the Jacobian of the corresponding transmission probability as

$$\begin{aligned} DT_{ij,kl} &= D_l|t_{ij,k}|^2 \\ &= D_l(\bar{t}_{ij,k}\, t_{ij,k}) \\ &= \overline{Dt}_{ij,kl}\, t_{ij,k} + \bar{t}_{ij,k}\, Dt_{ij,kl} \\ &= 2\,\mathrm{Re}(\bar{t}_{ij,k}\, Dt_{ij,kl}) \end{aligned} \tag{3.43}$$

where Re denotes the real part. The same formula applies to the reflection probabilities as well.

3.6 Verification

In this section we present some numerical results on the performance of the algorithms described in the previous chapter. To start with, numerical results on the direct FEM algorithm are shown. Next, comparison to the analytic solution is presented in a simple setting and Algorithm 3.2 is compared to brute force sensitivity calculations. Finally, effects of the magnetic field are investigated.

We consider a constant potential barrier as an example problem. The setup is the following. We choose our potential barrier to be of height $V_0 = 49$ in the spatial interval $[0, 1]$. The transmission and reflection probabilities (3.32) produced by Algorithm 3.1 are plotted in Figure 3.1. Since no magnetic field is present, only a single spin component is drawn. The black line represents the sum of T and R, and it is very close to one. This means that our solver preserves the probability well in this case. In this sense, the solution looks reasonable. We continue by comparing the transmission probability to the analytical solution.

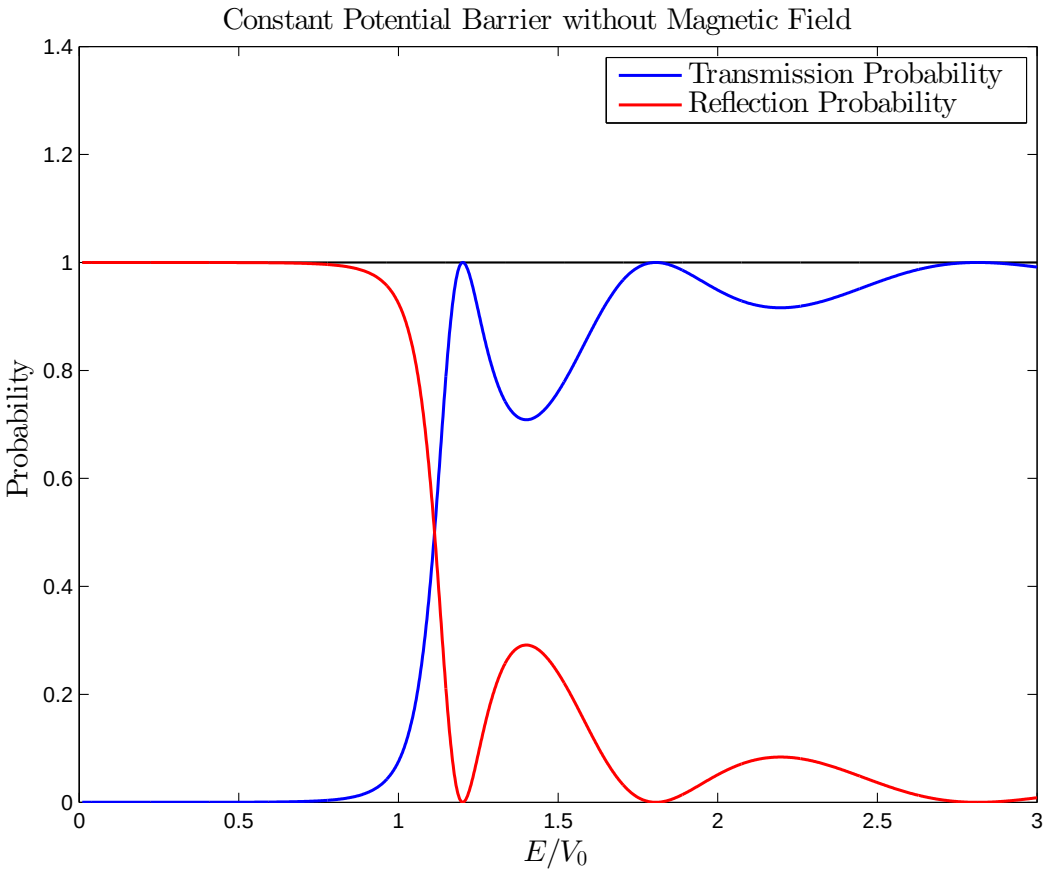

Figure 3.1: Transmission and reflection probabilities for a constant potential barrier.

For the constant potential barrier, the transmission probability is written as [13, ch. 7.7]

$$T(E) = \begin{cases} \dfrac{1}{1 + \frac{V_0^2}{4E(E-V_0)} \sin^2(C(E-V_0))}, & E > V_0, \\ \dfrac{1}{1 + \frac{V_0^2}{4E(V_0-E)} \sinh^2(C(V_0-E))}, & E < V_0, \\ \dfrac{1}{1 + \frac{CV_0}{4}}, & E = V, \end{cases}$$

where

$$C = \frac{2mL^2}{\hbar^2}.$$

The error between the two transmission probabilities is plotted in Figure 3.2 where we have used a finer energy mesh to get a smoother plot. The numerical error for Algorithm 3.1 depends on the number of elements in the FEM system. For this plot 99 elements were used. It seems that the relative error is larger when the energy is close to zero. Of course the transmission probability is also close to zero in that region and thus even small errors have a significant effect. After the energy exceeds the height of the barrier, the relative error decreases.

Since the derivation of the method for computing the Jacobian with respect to V was not a rigorous mathematical proof, we will verify it by numerical tests. We compare the solution of our derivate FEM program to a brute force method. The comparison is based on the following idea:

1. Solve derivatives, e.g., DT for the transmission probability using Algorithm 3.2 and equation (3.43).

2. Solve the brute force derivatives by

   (a) perturbing $V(x)$ at x_i by a small amount h,

   (b) solving the original FEM with Algorithm 3.1 at the perturbed potential,

   (c) computing
   $$D_i T_h = \frac{T_{\text{pertrubed}} - T}{h}$$

   (d) and repeating the process for all i to get the complete DT_h.

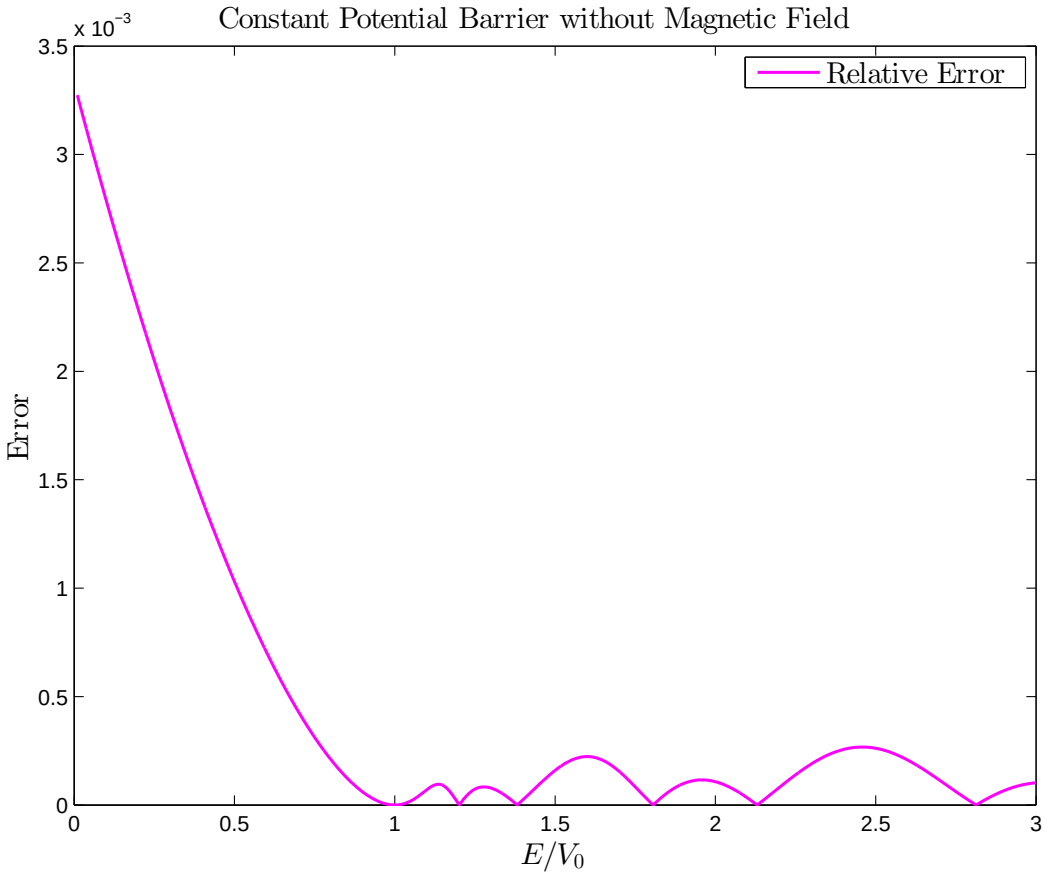

Figure 3.2: Relative numerical error in transmission probability.

In the following pictures we have used $h = 0.001$. The scattering problem is the same as considered earlier, i.e., the one involving a constant potential barrier. The derivative for the transmission probability produced by Algorithm 3.2 is drawn in Figure 3.3 and the relative error compared to the brute force derivative is shown in Figure 3.4. The x-axis indicates the perturbed point in $V(x)$. Looking at the derivative, we can see that the transmission probability is most sensitive to variations in the potential when E/V_0 is slightly over one. Also the changes in the potential are most ineffective near the end points of the interval $[0, 1]$. Comparing to Figure 3.1, one can roughly say that the transmission probability is sensitive at energy levels where it changes rapidly. From Figure 3.4 we also notice that the relative error between the two derivatives behaves well when the energy is small, that is, $E/V_0 < 1.0$, and even at higher energies the maximum error is only about 1 %. It is interesting to notice that some energy values are more prone to errors than others, e.g., when E/V_0 is approximately 1.2, 1.8 and 2.8.

Before concluding this chapter, a remark on the effect of magnetic fields is presented. As a sanity check, we consider a magnetic field with $B^1 = B^2 = 0$, which leads to

$$\boldsymbol{F}(x) = C \begin{bmatrix} E + B^3(x) - V(x) & 0 \\ 0 & E - B^3(x) - V(x) \end{bmatrix}$$

in equation (3.1). This means that the spin components do not mix but their interactions with the potential are different. Since a magnetic field is introduced, the diagonal of $\boldsymbol{F}$ has magnetic field term in addition to the potential $V(x)$. We now define the effective potentials for the two spins:

$$\begin{aligned} W_1(x) &= V(x) - B^3(x) \qquad \text{(up spin)}, \\ W_2(x) &= V(x) + B^3(x) \qquad \text{(down spin)}. \end{aligned}$$

This transforms $\boldsymbol{F}$ into

$$C \begin{bmatrix} E - W_1(x) & 0 \\ 0 & E - W_2(x) \end{bmatrix}$$

which looks like the original $\boldsymbol{F}$ without the magnetic field but with different potentials for the two spins. In other words, each spin component satisfies non-magnetic Schrödinger equation with the effective potentials W_1 and W_2. For a positive B^3, up spins should behave like down spins but in a weaker potential: For a constant potential barrier and constant magnetic field the up spin transmission probability should rise earlier than that of down spin. This is verified in Figure 3.5 where we used the same potential level $V_0 = 49$ as earlier and $B^3 = 20$.

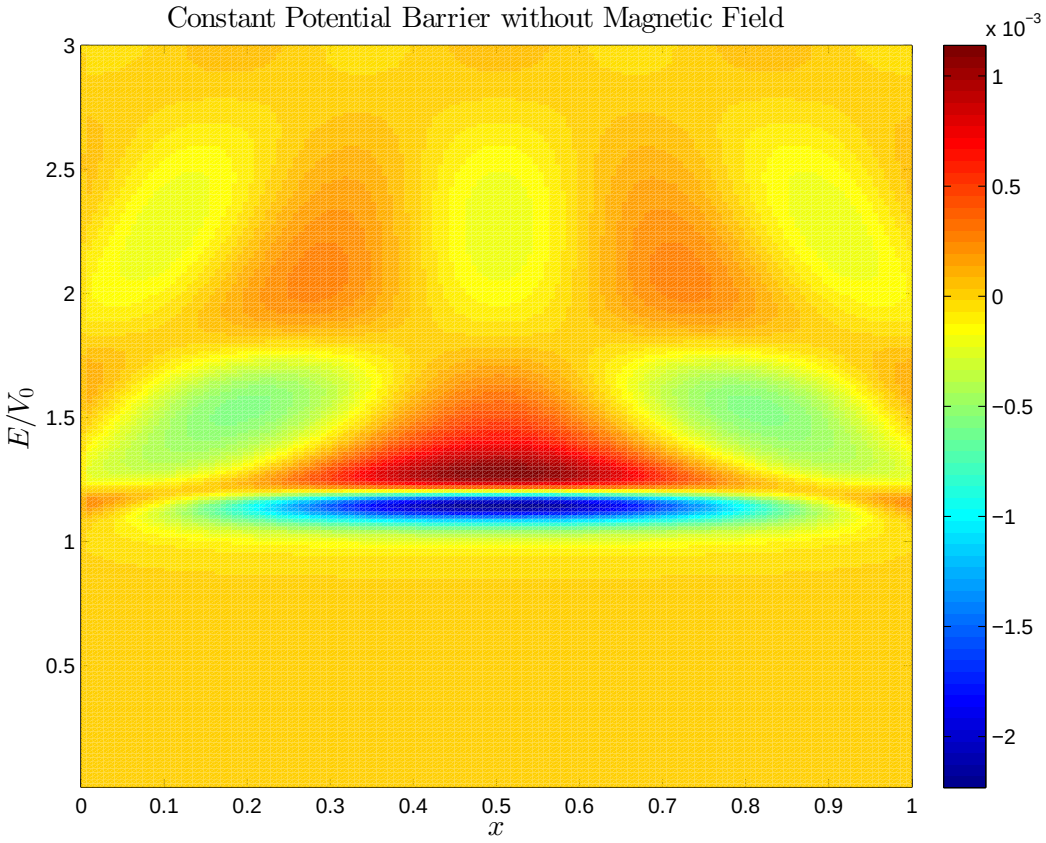

Figure 3.3: Transmission probability derivative.

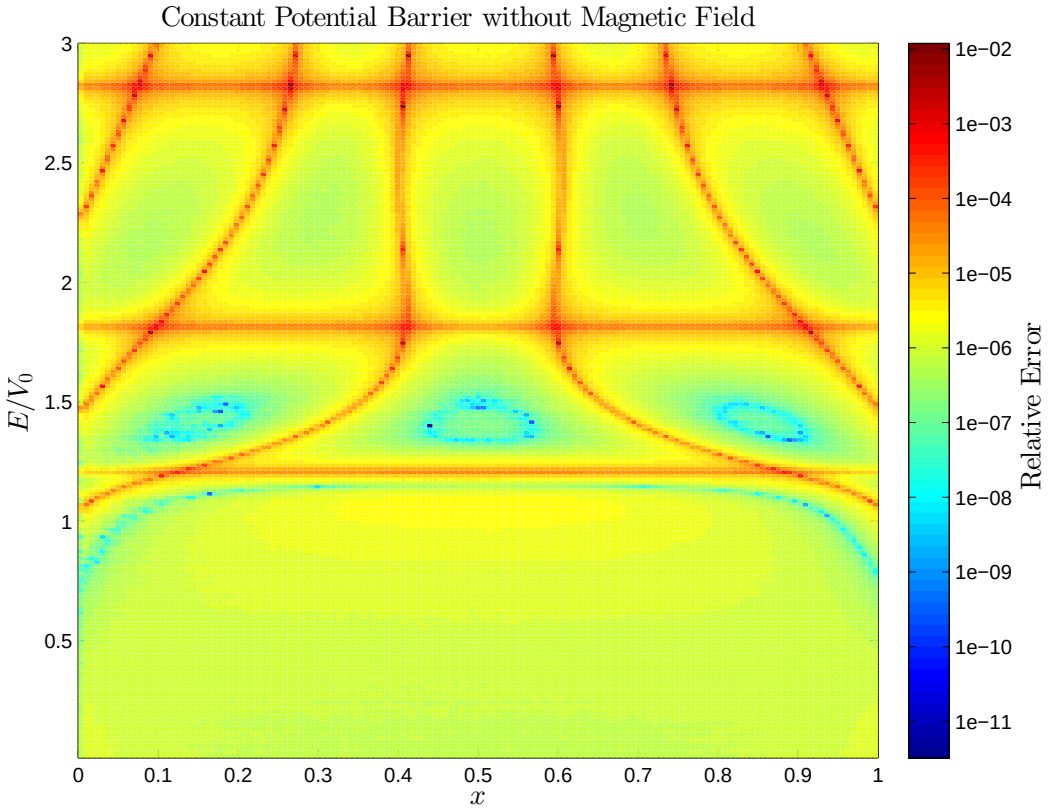

Figure 3.4: Relative error in transmission probability derivative.

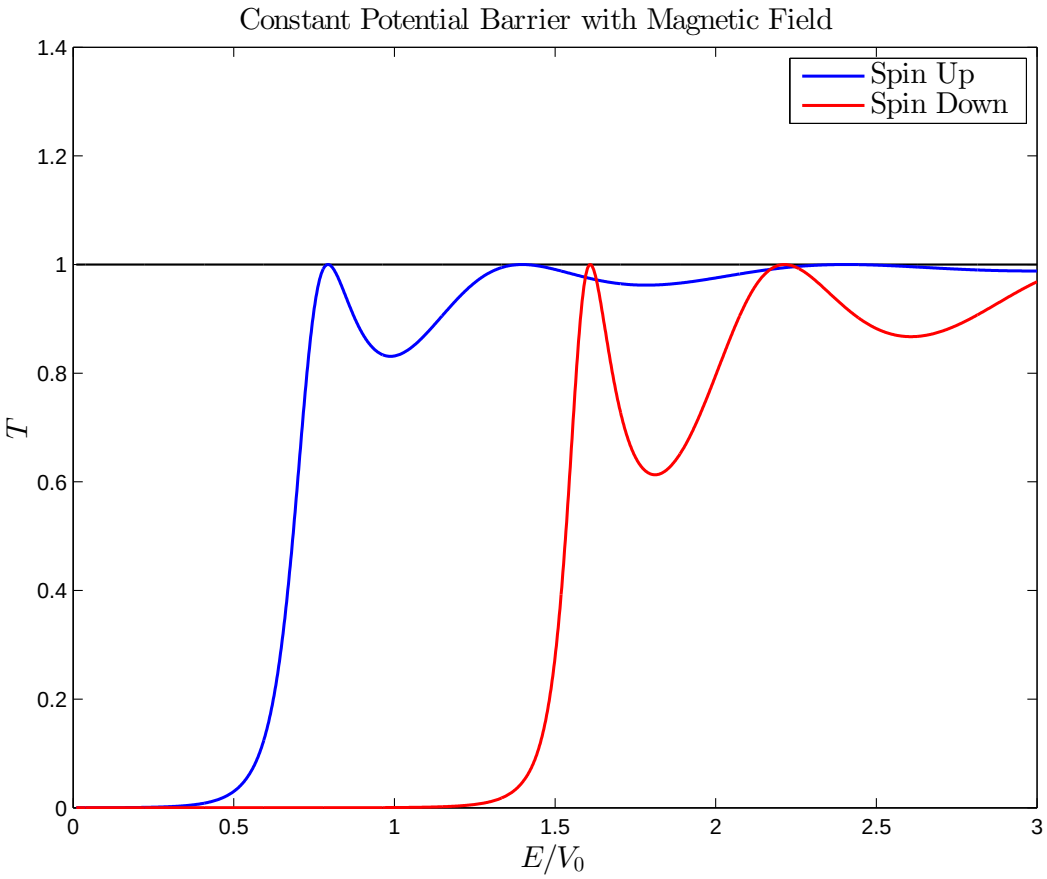

Figure 3.5: Transmission probability in magnetic field.

The other components B^1 and B^2 introduce mixing of the spins and thus their effects are more complicated. We investigate such mixing by adding B^2-directional component to the previously considered setup, that is, we set $V_0 = 49$, and $\boldsymbol{B} = [0, 15, 20]^T$. The transmission probabilities are presented in Figure 3.6. The upper plot corresponds to the transmission of up spin particles and the lower plot to that of down spin particles. The line drawn in black, shaded by blue and red areas, is the total transmission probability. The blue shading is the contribution of pure transmission, i.e., where no alteration of spin happens, and the red presents the effect of the change in spin. The blue line represents the transmission probability without change in spin and the red line that with spin change. For comparison, the transmission probabilities of the previous setting (Figure 3.5), are marked by a green line in both plots of Figure 3.6.

The behavior of up spin transmission is qualitatively similar for the two magnetic fields, that is, for $\boldsymbol{B} = [0, 0, 20]^T$ and $\boldsymbol{B} = [0, 15, 20]^T$. The addition of the B^2 component shifts the transmission probability curve slightly toward lower energies, but the total probability is in general a little lower when $E/V_0 < 1.5$. The largest part of the probability is constituted by the incoming up spin, which indicates only a little spin alteration. For the down spin, however, there is a remarkable difference. The mixing of the spins causes the down spin transmission to start much earlier: at $E/V_0 \approx 0.5$ compared to $E/V_0 \approx 1.5$ without B^2 component. In this interval, the transmission consist almost only of changing spin. The total transmission in the energy interval $[0.5, 1.5]$ is, however, less than 0.1. Above $E/V_0 \approx 1.1$, the total transmission probability has the same form as in the B^3-directional case, although there is a small shift in the direction of higher energies. It is interesting to notice that the effect of mixing is the same for both spins: T_{12} and T_{21} are identical. Thus, the difference between the spins is caused entirely by the B^3 component of the magnetic field.

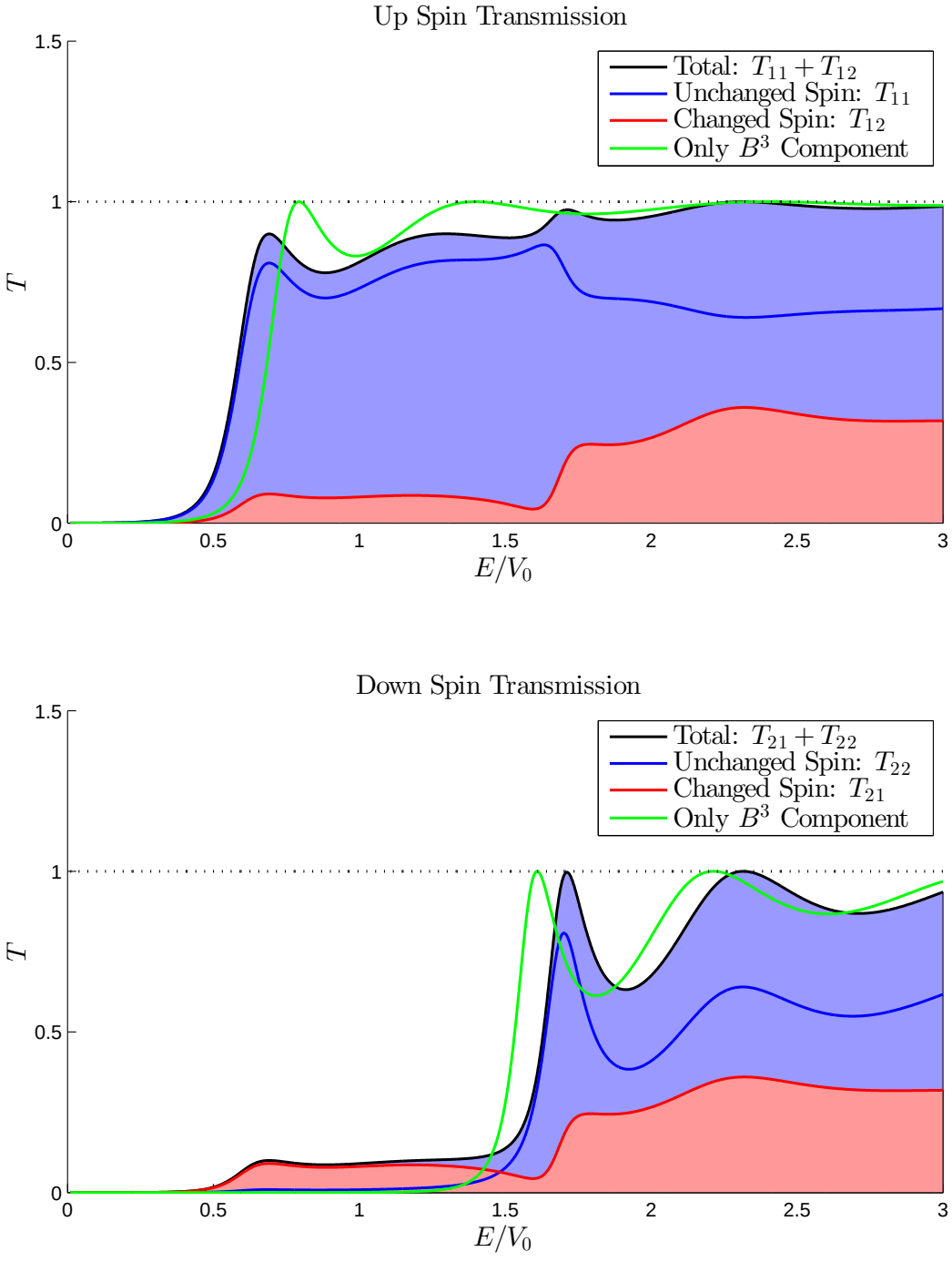

Figure 3.6: B^2-directional magnetic field induces mixing of the spins.

Algorithm 3.1 FEM Solver for Quantum Scattering

Create $\boldsymbol{K}$ (3.19), $\boldsymbol{V}_{12}$ and $\boldsymbol{V}_{21}$ (3.22)
Create containers $\mathcal{U}$, $\mathcal{T}$ and $\mathcal{R}$ for $\boldsymbol{u}_h$, $\boldsymbol{t}$ and $\boldsymbol{r}$ respectively
for $E \in \mathcal{E}$ **do**
  Update $\boldsymbol{V}_{11}$ and $\boldsymbol{V}_{22}$ (3.22)
  Build $\boldsymbol{V}$ (3.23)
  Include boundary conditions in $\boldsymbol{V}$ (3.25)
  $\boldsymbol{A} = \boldsymbol{K} + \boldsymbol{V}$
  $\boldsymbol{u}_h = \boldsymbol{0}$
  for $\boldsymbol{a}_+ \in \mathcal{A}$ **do**
    Update $\boldsymbol{f}$ (3.26)
    $\boldsymbol{v} = \boldsymbol{A}^{-1}\boldsymbol{f}$
    Compute $\boldsymbol{a}_+$ and $\boldsymbol{b}_-$ (3.31)
    Compute colunm of $\boldsymbol{t}$ and $\boldsymbol{r}$ (3.29), (3.30)
    $\boldsymbol{u}_h = \boldsymbol{u}_h + \boldsymbol{v}$
  end for
  Add to container: $\boldsymbol{u}_h \hookrightarrow \mathcal{U}$, $\boldsymbol{t} \hookrightarrow \mathcal{T}$ and $\boldsymbol{r} \hookrightarrow \mathcal{R}$
end for

Algorithm 3.2 FEM Solver with Derivative Included

Create $\boldsymbol{K}$ (3.19), $\boldsymbol{V}_{12}$ and $\boldsymbol{V}_{21}$ (3.22)
Create containers $\mathcal{U}$, $\mathcal{T}$, $\mathcal{R}$, $\mathcal{D}_{\mathcal{U}}$, $\mathcal{D}_{\mathcal{T}}$ and $\mathcal{D}_{\mathcal{R}}$, for $\boldsymbol{u}_h$, $\boldsymbol{t}$, $\boldsymbol{r}$, $D\boldsymbol{u}_h$, $D\boldsymbol{t}$ and $D\boldsymbol{r}$ respectively
for $E \in \mathcal{E}$ **do**
  Update $\boldsymbol{V}_{11}$ and $\boldsymbol{V}_{22}$ (3.22)
  Build $\boldsymbol{V}$ (3.23)
  Include boundary conditions in $\boldsymbol{V}$ (3.25)
  $\boldsymbol{A} = \boldsymbol{K} + \boldsymbol{V}$
  $\boldsymbol{u}_h = \boldsymbol{0}$
  for $\boldsymbol{a}_+ \in \mathcal{A}$ **do**
    Update $\boldsymbol{f}$ (3.26)
    $\boldsymbol{v} = \boldsymbol{A}^{-1}\boldsymbol{f}$
    Compute $\boldsymbol{a}_+$ and $\boldsymbol{b}_-$ (3.31)
    Compute column of $\boldsymbol{t}$ and $\boldsymbol{r}$ (3.29), (3.30)
    $\boldsymbol{u}_h = \boldsymbol{u}_h + \boldsymbol{v}$
  end for
  Add to container: $\boldsymbol{u}_h \hookrightarrow \mathcal{U}$, $\boldsymbol{t} \hookrightarrow \mathcal{T}$ and $\boldsymbol{r} \hookrightarrow \mathcal{R}$
  Allocate storage for $D\boldsymbol{u}$, $D\boldsymbol{t}$ and $D\boldsymbol{r}$
  for all Grid points x_l **do**
    Build $\boldsymbol{y}$ (3.40) (3.39) (3.21)
    $D_l\boldsymbol{u}_h = \boldsymbol{0}$
    for $\boldsymbol{a}_+ \in \mathcal{A}$ **do**
      $D_l\boldsymbol{v} = \boldsymbol{A}^{-1}\boldsymbol{y}$
      Compute $D_l\boldsymbol{a}_+$ and $D_l\boldsymbol{b}_-$ (3.41)
      Compute components of $D_l\boldsymbol{t}$ and $D_l\boldsymbol{r}$ (3.42)
      $D_l\boldsymbol{u}_h = D_l\boldsymbol{u}_h + D_l\boldsymbol{v}$
    end for
    Include $D_l\boldsymbol{u}$, $D_l\boldsymbol{t}$ and $D_l\boldsymbol{r}$ in $D\boldsymbol{u}$, $D\boldsymbol{t}$ and $D\boldsymbol{r}$
  end for
  Add to container: $D\boldsymbol{u}_h \hookrightarrow \mathcal{D}_{\mathcal{U}}$, $D\boldsymbol{t} \hookrightarrow \mathcal{D}_{\mathcal{T}}$ and $D\boldsymbol{r} \hookrightarrow \mathcal{D}_{\mathcal{R}}$
end for

Chapter 4

Inverse Problem

This chapter builds on the previous one and is based on the text book by Kaipio and Somersalo [11]. The goal is to write an algorithm for inferring the potential $V(x)$ from the transmission probabilities; for more information on inverse scattering problems, see [5] and the references therein. As we are considering both up and down spins, we could in principle use all four components of

$$\boldsymbol{T}(E) = \begin{bmatrix} T_{11}(E) & T_{12}(E) \\ T_{21}(E) & T_{22}(E) \end{bmatrix}, \tag{4.1}$$

but for simplicity we consider only one of the components for now. The resulting program will be described for all components in Section 4.3. Moreover, there is no need to consider reflection probabilities since they do not add any information because

$$R = 1 - T.$$

We can interpret our forward solver as a non-linear transformation

$$A : V(x) \mapsto T(E)$$

for discretized potentials V and resulting transmission probabilities T. In the corresponding inverse problem, we want to find some $V(x)$ that produces a given $T(E)$. If our equation was linear, we would be able to represent the solver as a linear transformation, say matrix $\boldsymbol{A}$, and solve

$$\boldsymbol{A}\boldsymbol{x}_V = \boldsymbol{y}_T$$

for $\boldsymbol{x}_V$ in some regularized manner. In this formula $\boldsymbol{x}_V$ and $\boldsymbol{y}_T$ represent the discretized versions of $V(x)$ and $T(E)$ respectively. Very loosely speaking, regularization means not solving the system exactly. A reason for this is that

non-physical properties for $V(x)$ may arise during the solving procedure due to instability. In addition, if $T(E)$ is measured and thus contains unavoidable inaccuracies, the reconstructed potential function is typically ruined without proper regularization. This setup creates a need for subtle tuning of the inverse solver.

To make the situation worse, the scattering problem (3.1) is non-linear with respect to $V(x)$. This means that a straightforward application of the above method is out of question. However, since we can compute the sensitivity of the solver with respect to $V(x)$, we can consider our solver being linear when the change in the potential is small. This allows us to iteratively find a possible solution to the non-linear problem by taking small successive steps. This approach is described in detail in the next sections.

4.1 Regularization of Inverse Problems

In inverse problems the aim is to deduce the unknown $\boldsymbol{x} \in \mathbb{R}^n$ from the consequences, or measurements, $\boldsymbol{y} \in \mathbb{R}^m$. We consider A as an operator which maps the data to the measurements. Given $\boldsymbol{y}$, we want to find as accurate $\boldsymbol{x}$ as possible. We now write our scattering problem in an abstract form

$$A(\boldsymbol{x}) = \boldsymbol{y}$$

and return to explicitly using the potential and transmission probabilities in Section 4.3. In order to find $\boldsymbol{y}$, the straightforward approach is to minimize

$$\|A(\boldsymbol{x}) - \boldsymbol{y}\|^2.$$

For a finite-dimensional linear A, this is just an ordinary least squares problem which has the solution

$$\boldsymbol{x} = (A^T A)^{-1} A^T \boldsymbol{y},$$

assuming $A^T A$ is invertible.

Due to instability, the least squares method is not a good way to proceed and thus some modifications are needed. The leading idea is to include some prior knowledge on $\boldsymbol{x}$ in the minimization process. We might want $\boldsymbol{x}$ to be smooth or small or to carry some other desired properties. To this end, we introduce a regularization term $G : \mathbb{R}^n \to \mathbb{R}_+$ and solve

$$\underset{\boldsymbol{x} \in \mathbb{R}^n}{\arg\min} \ \|A(\boldsymbol{x}) - \boldsymbol{y}\|^2 + \delta G(\boldsymbol{x}). \tag{4.2}$$

This is called Tikhonov regularization [9]. The parameter $\delta > 0$ is introduced to control the amount of regularization. The simplest choice would be

$$G(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$$

which leads to the minimization of

$$\|A(\boldsymbol{x}) - \boldsymbol{y}\|^2 + \delta\|\boldsymbol{x}\|^2.$$

In this case large norms of $\boldsymbol{x}$ are penalized and the amount of penalty is controlled by δ. The choice of G and δ depends very much on the problem at hand.

If the data contains noise, some technique to tune the regularization parameter δ is needed. One commonly used method is called the Morozov discrepancy principle [9]. As a motivation, consider the noise model

$$\boldsymbol{y} = \boldsymbol{y}' + \boldsymbol{e}$$

where $\boldsymbol{y}$ is the available measurement, $\boldsymbol{y}'$ is the exact data and $\boldsymbol{e}$ is a realization of some random variable modeling the noise. In principle, we would like our reconstruction $\boldsymbol{x}_\delta$ to minimize

$$\|A(\boldsymbol{x}_\delta) - \boldsymbol{y}\|.$$

However, since there is noise in $\boldsymbol{y}$, even if the reconstruction was perfect, i.e., $A(\boldsymbol{x}_\delta) = \boldsymbol{y}$, we would have the residual

$$\|A(\boldsymbol{x}_\delta) - \boldsymbol{y}'\| = \|\boldsymbol{e}\|.$$

Thus, when tuning the δ parameter, we need to pay attention to the values of

$$f(\delta) = \|A(\boldsymbol{x}_\delta) - \boldsymbol{y}\|.$$

Letting $f(\delta)$ be smaller than the norm of the noise would roughly mean that we are trying to fit $\boldsymbol{x}_d$ to the noise. Even though the realization of the noise vector is usually unknown, we may assume to know some estimate $\varepsilon \approx \|\boldsymbol{e}\|$. The idea of the Morozov discrepancy principle is to decrease δ until

$$f(\delta) \leq \varepsilon.$$

In a finite dimensional setting, a possible estimate for the norm $\|\boldsymbol{e}\|$ is, for example,

$$\varepsilon = \sigma\sqrt{m} \tag{4.3}$$

where σ is the estimated standard deviation of the mean-free noise and m is the dimension of the noise vector.

4.2 Gauss–Newton Method

We continue by describing a method for solving non-linear inverse problems in connection with Tikhonov regularization. The procedure relies on the ability to linearize the equation (4.2). Since the regularization term must also be linearized, we employ for simplicity

$$G(\boldsymbol{x}) = \|\boldsymbol{L}\boldsymbol{x}\|^2 \tag{4.4}$$

where $\boldsymbol{L}$ is some suitably chosen matrix. Thus we need to minimize

$$F(\boldsymbol{x}) = \|A(\boldsymbol{x}) - \boldsymbol{y}\|^2 + \delta\|\boldsymbol{L}\boldsymbol{x}\|^2, \quad \delta > 0,$$

where $A(x)$ is a non-linear map $A : \mathbb{R}^n \to \mathbb{R}^m$, and the regularization matrix $L \in \mathbb{R}^{r\times n}$ for some $r \in \mathbb{N}$. In Section 4.3, we will see that this is the form we get by discretizing our inverse scattering problem.

Next, we derive the Gauss–Newton method for minimizing F. Details of the algorithm can be found, e.g., in [14]. The basic idea behind the procedure is to get a new approximate solution $\boldsymbol{x}'$ by taking a small step $\boldsymbol{p}$,

$$\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{p}.$$

The step $\boldsymbol{p}$ is in principle solved from the following equation:

$$\mathrm{Hess}(F)(\boldsymbol{x})\,\boldsymbol{p} = -\nabla F(\boldsymbol{x}) \tag{4.5}$$

where the Hessian, $\mathrm{Hess}(F)$, and the gradient, ∇F, are evaluated at the previous value $\boldsymbol{x}$. Observe that (4.5) follows by linearizing the necessary condition for the minimizer of F around $\boldsymbol{x}$. In the Gauss–Newton method the Hessian is approximated in order to achieve a numerically less intensive scheme: one wants to avoid computing the second order derivatives in the Hessian.

The gradient of F is

$$\begin{aligned}\nabla F(\boldsymbol{x}) &= \nabla\big((A(\boldsymbol{x}) - \boldsymbol{y})^T(A(\boldsymbol{x}) - \boldsymbol{y}) + \delta(\boldsymbol{L}\boldsymbol{x})^T\boldsymbol{L}\boldsymbol{x}\big) \\ &= 2\left(\boldsymbol{J}_A(\boldsymbol{x})^T(A(\boldsymbol{x}) - \boldsymbol{y}) + \delta\boldsymbol{L}^T\boldsymbol{L}\boldsymbol{x}\right)\end{aligned} \tag{4.6}$$

where $\boldsymbol{J}_A$ is the Jacobian of A

$$[\boldsymbol{J}_A]_{ij} = \frac{\partial A_i}{\partial x_j}$$

when denoting the components of $A(\boldsymbol{x})$ by A_i. Note that we use the 'real' transpose instead of the conjugate transpose since A and $\boldsymbol{L}$ can be

considered real in our scattering setting. To obtain the Hessian, we need to differentiate for a second time

$$\begin{aligned}\left[\mathrm{Hess}(F)(\boldsymbol{x})\right]_{ij} &= \frac{\partial}{\partial x_j}\left(2\left(\sum_{k=1}^{m}\frac{\partial A_k}{\partial x_i}(A_k - y_k) + \delta\sum_{l=1}^{n}\sum_{s=1}^{r} L_{si}L_{sl}\,x_l\right)\right)\\ &= 2\Bigg(\sum_{k=1}^{m}\left(\frac{\partial^2 A_k}{\partial x_i \partial x_j}(A_k - y_k) + \frac{\partial A_k}{\partial x_i}\frac{\partial A_k}{\partial x_j}\right)\\ &\qquad + \delta\sum_{l=1}^{n}\sum_{s=1}^{r} L_{si}L_{sl}\,\frac{\partial x_l}{\partial x_j}\Bigg).\end{aligned}$$

Excluding the second order derivatives and noting that

$$\frac{\partial x_k}{\partial x_j} = I_{lj} = \begin{cases}1, & l = j,\\ 0, & l \neq j,\end{cases}$$

the approximate Hessian is written as

$$\begin{aligned}\left[\mathrm{Hess}(F)(\boldsymbol{x})\right]_{ij} &\approx 2\left(\sum_{k=1}^{m}\frac{\partial A_k}{\partial x_i}\frac{\partial A_k}{\partial x_j} + \delta\sum_{l=1}^{n}\sum_{s=1}^{r} L_{si}L_{sl}I_{lj}\right)\\ &= 2\left(\sum_{k=1}^{m}\frac{\partial A_k}{\partial x_i}\frac{\partial A_k}{\partial x_j} + \delta\sum_{s=1}^{r} L_{si}L_{sj}\right),\end{aligned}$$

or in matrix form

$$\mathrm{Hess}(F)(\boldsymbol{x}) \approx 2\left(\boldsymbol{J}_A(\boldsymbol{x})^T\boldsymbol{J}_A(\boldsymbol{x}) + \delta\boldsymbol{L}^T\boldsymbol{L}\right). \tag{4.7}$$

The Gauss–Newton method will thus solve the equation (4.5) with the approximate Hessian given by (4.7) and the gradient by (4.6). Writing (4.5) with the help of $\boldsymbol{x}$ instead of $\boldsymbol{p}$ leads to

$$\mathrm{Hess}(F)(x)\,\boldsymbol{x}' = -\nabla F(x) + \mathrm{Hess}(F)(x)\,\boldsymbol{x}.$$

Plugging in the approximate Hessian (4.7) and the gradient (4.6) gives

$$\begin{aligned}2\left(\boldsymbol{J}_A(\boldsymbol{x})^T\boldsymbol{J}_A(\boldsymbol{x}) + \delta\boldsymbol{L}^T\boldsymbol{L}\right)\boldsymbol{x}' = &-2\left(\boldsymbol{J}_A(\boldsymbol{x})^T(A(\boldsymbol{x}) - \boldsymbol{y}) + \delta\boldsymbol{L}^T\boldsymbol{L}\boldsymbol{x}\right)\\ &+ 2\left(\boldsymbol{J}_A(\boldsymbol{x})^T\boldsymbol{J}_A(\boldsymbol{x}) + \delta\boldsymbol{L}^T\boldsymbol{L}\right)\boldsymbol{x},\end{aligned}$$

which simplifies to

$$\left(\boldsymbol{J}_A(\boldsymbol{x})^T\boldsymbol{J}_A(\boldsymbol{x}) + \delta\boldsymbol{L}^T\boldsymbol{L}\right)\boldsymbol{x}' = \boldsymbol{J}_A(\boldsymbol{x})^T(\boldsymbol{y} - A(\boldsymbol{x}) + \boldsymbol{J}_A(\boldsymbol{x})\,\boldsymbol{x}).$$

In the Gauss–Newton method this equation is repeatedly solved to obtain a new, hopefully more accurate, solutions to the minimization problem:

$$\left(\boldsymbol{J}_A^{(k)^T}\boldsymbol{J}_A^{(k)} + \delta \boldsymbol{L}^T\boldsymbol{L}\right)\boldsymbol{x}^{(k+1)} = \boldsymbol{J}_A^{(k)^T}\left(\boldsymbol{y} - A(\boldsymbol{x}^{(k)}) + \boldsymbol{J}_A^{(k)}\boldsymbol{x}^{(k)}\right) \tag{4.8}$$

where k denotes the iteration round. The convergence of the Gauss–Newton method depends on the initial guess. If the starting point is good, the convergence is fast, and in optimal case quadratic [14, ch. 10.2]. Moreover, when using the method for inverse problems with regularization, the parameter δ plays a key role in tuning the performance.

4.3 Implementation

In this section we describe an algorithm for reconstructing the potential function $V(x)$ when knowing the transmission probabilities $\boldsymbol{T}(E)$ (4.1). Remember that our numerical algorithm only computes the amplitudes $\boldsymbol{t}$ and their derivatives. The probabilities are then obtained via (3.32), and the amplitude derivatives can be transformed to probability derivatives using the chain rule, cf. (3.43). Thus, the toolbox assumed for the inverse algorithm is having a program which generates $\boldsymbol{T}$ and $D\boldsymbol{T}$ from the discretized potential function $V(x)$. In essence, Algorithm 3.2 complemented with the above transformations describes such a program.

The main task of this section is solving the discretized version of the minimization problem

$$\underset{V(x)}{\arg\min}\ \|A(V(x)) - \boldsymbol{T}(E)\|^2 + \delta\|L\,V(x)\|^2$$

which is rewritten version of (4.2) with the linear regularization term

$$G(V) = \|L\,V(x)\|^2.$$

By discretizing, we end up solving

$$\underset{\boldsymbol{v}}{\arg\min}\ \|A(\boldsymbol{v}) - \boldsymbol{\tau}\|^2 + \delta\|\boldsymbol{L}\boldsymbol{v}\|^2, \tag{4.9}$$

where $A : \boldsymbol{v} \mapsto \boldsymbol{\tau}$ is related to the discrete forward solver described in Chapter 3 and $\boldsymbol{L}$ is some regularization matrix. Here $\boldsymbol{v}$ and $\boldsymbol{\tau}$ denote the discretized $V(x)$ and $\boldsymbol{T}(E)$, respectively. These variables arise from the design choice to include all components of $\boldsymbol{T}$ in the inversion. In order to write the minimization problem explicitly and apply Gauss–Newton, we need to fine-tune the solver described in Chapter 3.

First, $\boldsymbol{T}(E)$ needs to be presented in a vector form. Secondly, the Jacobian of A, which is used in (4.9), needs to be created from $D\boldsymbol{T}$. We start by vectorizing $\boldsymbol{T}$. Let

$$\boldsymbol{\tau}_{ij} = \begin{bmatrix} T_{ij}(E_1) \\ T_{ij}(E_2) \\ \vdots \\ T_{ij}(E_{N_e}) \end{bmatrix} \tag{4.10}$$

where N_e denotes the number of computed energy levels. Subsequently, we get the vector form of $\boldsymbol{T}$ by stacking the components of $\boldsymbol{T}$ as follows:

$$\boldsymbol{\tau} = \begin{bmatrix} \boldsymbol{\tau}_{11} \\ \boldsymbol{\tau}_{21} \\ \boldsymbol{\tau}_{12} \\ \boldsymbol{\tau}_{22} \end{bmatrix}. \tag{4.11}$$

Next we need to form a compatible Jacobian based on $D\boldsymbol{T}$. In a sense, $D\boldsymbol{T}$ is a four-dimensional array since it contains Jacobians for each T_{ij}. We denote

$$DT_{ijkl} = \frac{\partial T_{ij}(E_k)}{\partial V_l}$$

which means that the transmission probability derivative of T_{ij} is evaluated at E_k when $V(x_l)$ is perturbed. Thus the Jacobian we are looking for is just

$$\boldsymbol{J}_{ij} = D\boldsymbol{T}_{ij} \tag{4.12}$$

and the Jacobian for all components is

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}_{11} \\ \boldsymbol{J}_{21} \\ \boldsymbol{J}_{12} \\ \boldsymbol{J}_{22} \end{bmatrix}. \tag{4.13}$$

We discretize $V(x)$ in the standard way:

$$\boldsymbol{v} = \begin{bmatrix} V(x_1) \\ V(x_2) \\ \vdots \\ V(x_N) \end{bmatrix}.$$

The mesh points $x_1, \ldots, x_N$ must be uniformly spaced because the Gauss–Newton iteration uses the forward solver which requires a uniform mesh.

The details of the implementation are found in Algorithm 4.1 where we have presented the code for the above introduced vectorization. Note that one could use other vectorization as well; only the modification of $\boldsymbol{\tau}$ and the according formation of the Jacobian need to be done. For example, if we were interested in total spin up and down transmissions instead of considering all four components, we would have

$$\boldsymbol{\tau} = \begin{bmatrix} \boldsymbol{\tau}_{11} + \boldsymbol{\tau}_{12} \\ \boldsymbol{\tau}_{21} + \boldsymbol{\tau}_{22} \end{bmatrix}$$

and

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}_{11} + \boldsymbol{J}_{12} \\ \boldsymbol{J}_{21} + \boldsymbol{J}_{22} \end{bmatrix}$$

where $\boldsymbol{\tau}_{ij}$ and $\boldsymbol{J}_{ij}$ are from (4.10) and (4.12), respectively. This is exactly what we are going to do in the next section when we investigate how the reconstruction algorithm fares with noisy data.

4.4 Numerical Examples

In this section we present some reconstructions of potentials by Algorithm 4.1. The images are created by first solving the forward problem with a given potential, then adding Gaussian noise with zero mean and standard deviation 0.01 to the components of the data vector, and finally utilizing the inverse solver. Since we add artificial measurement noise, the natural way to decide the degree of regularization is to use the Morozov discrepancy principle to determine the Gauss–Newton stopping parameter ε, cf. Algorithm 4.1. The formula for computing ε can be found in (4.3).

We start by inverting the constant potential barrier considered in Section 3.6. We also include a magnetic field. The potential and the magnetic field are the same as in Figure 3.5 previously, i.e., $V_0 = 49$ and $B = [0, 0, 20]^T$. The reconstructed potential is shown in Figure 4.1, where the original, correct potential is plotted in green and the reconstruction in blue. The red dashed line is the initial guess used in the Gauss–Newton algorithm. The transmission probabilities computed with the reconstruction can be found in Figure 4.2, along with the noisy initial data drawn in black. Naturally, the inverted potential is not an exact copy of the target potential but has some oscillations. The transmission probabilities corresponding to the reconstructed potential are however very close to the noisy data, which is a manifestation of the illposedness of the consider inverse scattering problem: large changes in the potential may alter the data only a little. The used

regularization term was

$$G(V) = \|V(x)\|^2$$

which corresponds to $\boldsymbol{L}$ being the identity matrix in (4.4). The algorithm is very sensitive to the choice of the regularization parameter δ. In Figure 4.3, one can see the reconstruction with a slightly smaller δ. The solution has disordered, non-physical behavior; in this case the Gauss–Newton iteration does not even converge.

Correct tuning of the regularization parameter and the choice of $G(x)$ enable the algorithm to produce a good reconstruction. To verify this claim, we examine the use of another penalty term. To prefer smooth solutions, one can require the second derivative of the potential function to be small. To this end, we choose

$$G(V) = \|V''(x)\|^2.$$

which needs to be discretized to include it in the reconstruction algorithm. The difference approximation for the second derivative with homogeneous Dirichlet boundary conditions is given by the following matrix

$$\frac{1}{h^2}\begin{bmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix},$$

where h is the mesh parameter which we exclude here since it can be included in the regularization parameter δ. In consequence, our regularization matrix is

$$\boldsymbol{L} = \begin{bmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix}. \tag{4.14}$$

This smoothness prior is used in the next example.

The considered potential has two bumps instead of one:

$$V(x) = V_{\max} \sin^2(2\pi x), \quad x \in [0, 1].$$

Again we solve the direct problem with this potential, add about 1% of noise and invert. The reconstructed potential is presented in Figure 4.4 and

Algorithm 4.1 Gauss–Newton Solver

Input: Possibly noisy data $\boldsymbol{\tau}$
Parameters: Initial guess $\boldsymbol{v}^{(0)}$, regularization terms δ and $\boldsymbol{L}$
Iteration stopping conditions: ε and maximum iterations N_i

for $0 \leq k < N_i$ **do**
  Solve $\boldsymbol{\tau}^{(k)}$ and $\boldsymbol{J}^{(k)}$ for $\boldsymbol{v}^{(k)}$ with the Algorithm 3.2 and the Equations (4.10), (4.11), (4.12) and (4.13)
  if $\|\boldsymbol{\tau}^{(k)} - \boldsymbol{\tau}\| < \varepsilon$ **then**
    return $\boldsymbol{v}^{(k)}$
  end if
  Solve $\left(\boldsymbol{J}^{(k)^*}\boldsymbol{J}^{(k)} + \delta \boldsymbol{L}^*\boldsymbol{L}\right)\boldsymbol{v}^{(k+1)} = \boldsymbol{J}^{(k)^*}\left(\boldsymbol{\tau} - \boldsymbol{\tau}^{(k)} + \boldsymbol{J}^{(k)}\boldsymbol{x}^{(k)}\right)$
end for
return $\boldsymbol{v}^{(k+1)}$

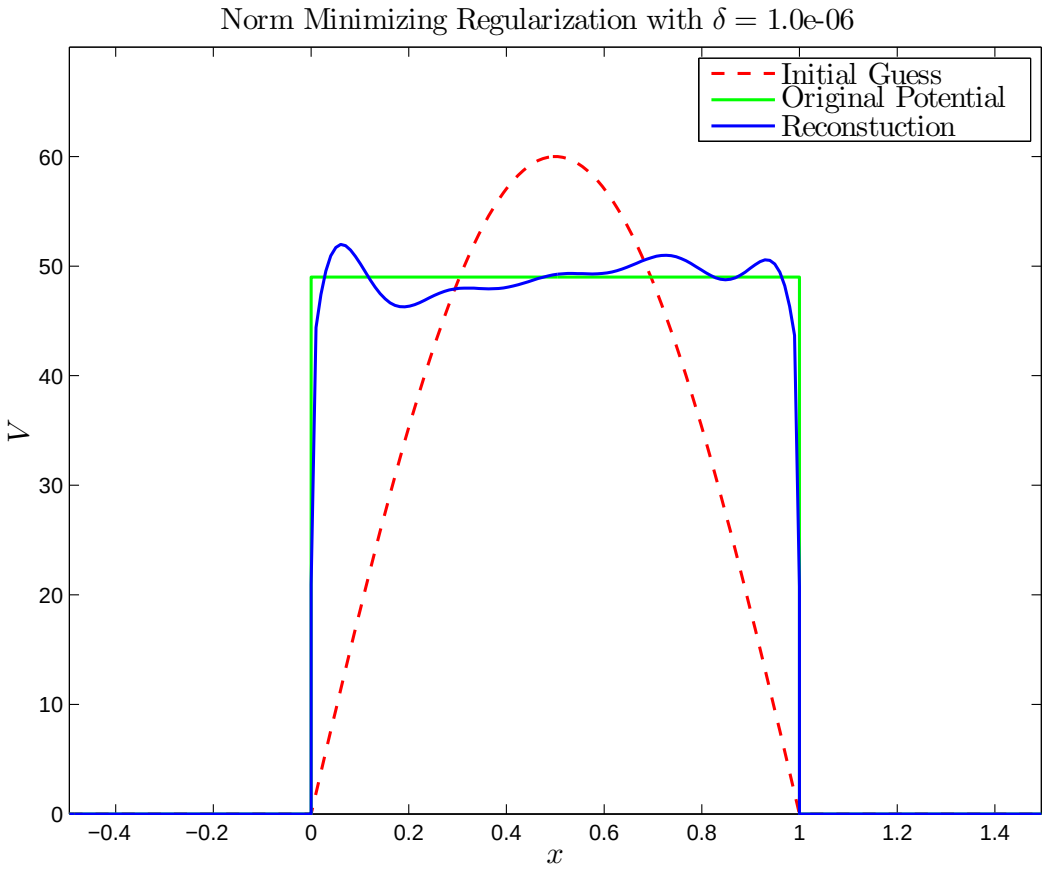

Figure 4.1: Reconstruction of a constant potential barrier.

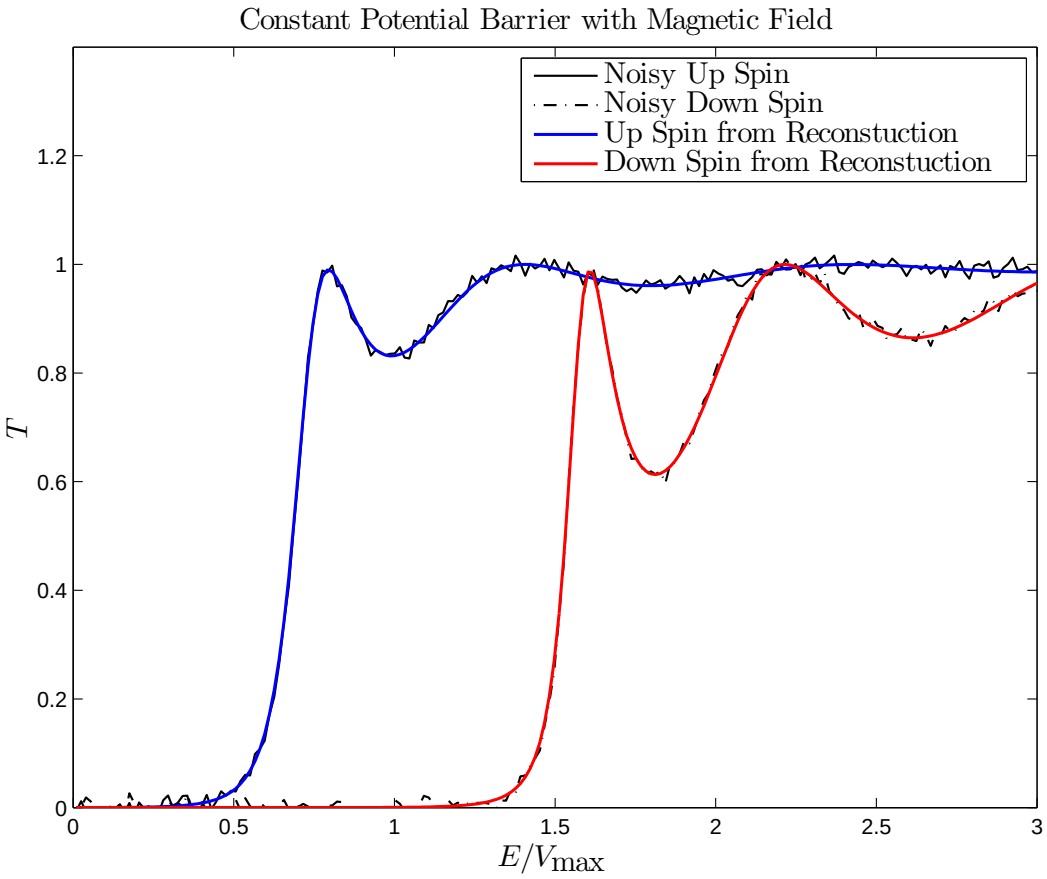

Figure 4.2: Transmission probability of the reconstructed constant potential barrier.

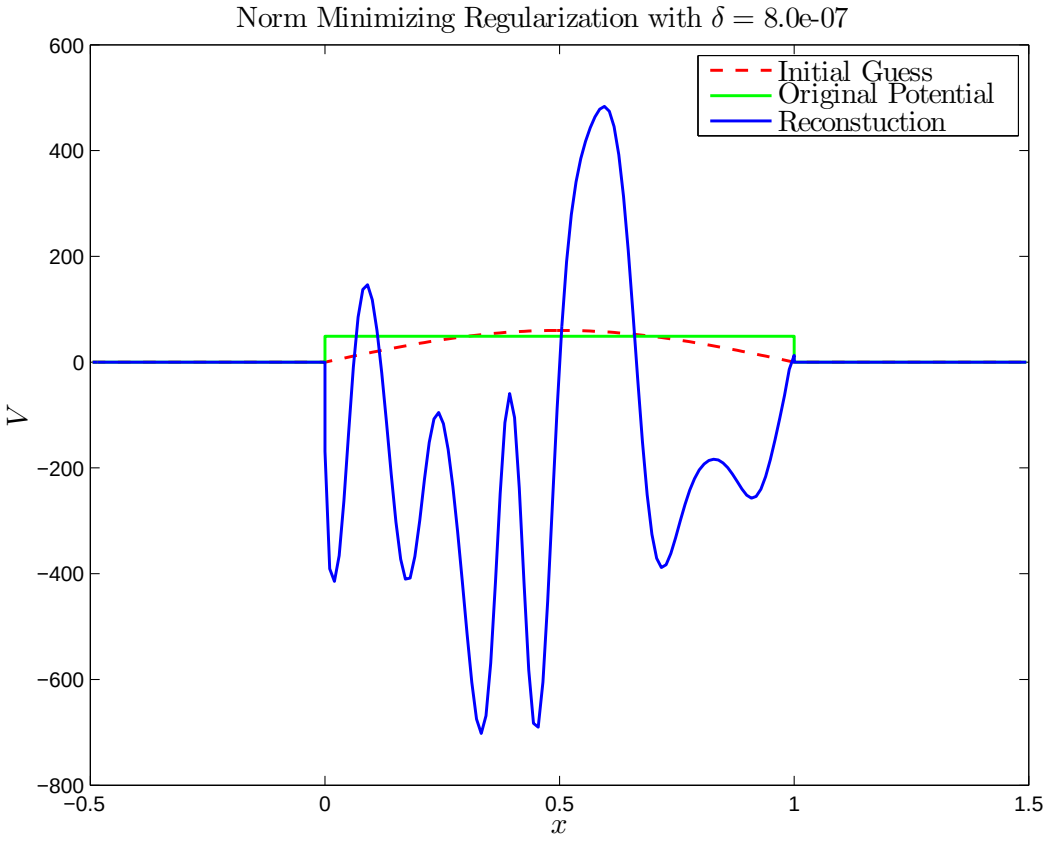

Figure 4.3: Reconstruction is highly dependent on the choice of the regularization parameter.

the corresponding transmission probabilities in Figure 4.5. We have chosen $V_{\max} = 30$, and the magnetic field remains the same as in the previous example. With a correct choice of δ, we are able to obtain good results. Compared to the constant potential, the reconstruction is more accurate and the reproduction of the transmission probabilities works equally well.

In the previous tests we have used additive Gaussian noise with standard deviation 0.01. In order to study the effect of noise on the reconstruction, we return to the constant potential barrier case. We use the same setup as in Figures 4.1 and 4.2, that is, $V_0 = 49$ and $\boldsymbol{B} = [0, 0, 20]^T$. In Figure 4.6, the reconstructions of the potential are plotted with the standard deviations $\sigma = 0.001, 0.01, 0.1$ for the noise. With the largest σ, the reconstructed potential has clearly more oscillations than for the lower noise levels. However, the difference between the reconstructions with $\sigma = 0.01$ and $\sigma = 0.001$ is small. The regularization parameter δ is the same as in Figure 4.1.

In the presented examples, we have used norm minimizing regularization for all potentials, with the exception of the 'double-bump' case. To complete this section, we now examine how a smoothness prior affects the constant barrier reconstruction. The comparison is made against the norm minimizing regularization with the same amount of noise, $\sigma = 0.01$. The smoothness matrix is the same as before, i.e., (4.14), but the parameter δ was tuned by hand to obtain a good reconstruction. The reconstructed potentials are presented in Figure 4.7. Unsurprisingly, the smoothness regularization implies less oscillations. More importantly, the values near the end points of the interval $[0, 1]$ are lower since the regularization assumes zero values at the boundary, i.e., at 0 and 1. Thus non-zero potential near the end points is penalized. If one decreases δ, more oscillations will occur and the reconstruction is more similar to the norm minimizing one. However, sharp edges at the end points cannot be obtained with the employed smoothness prior.

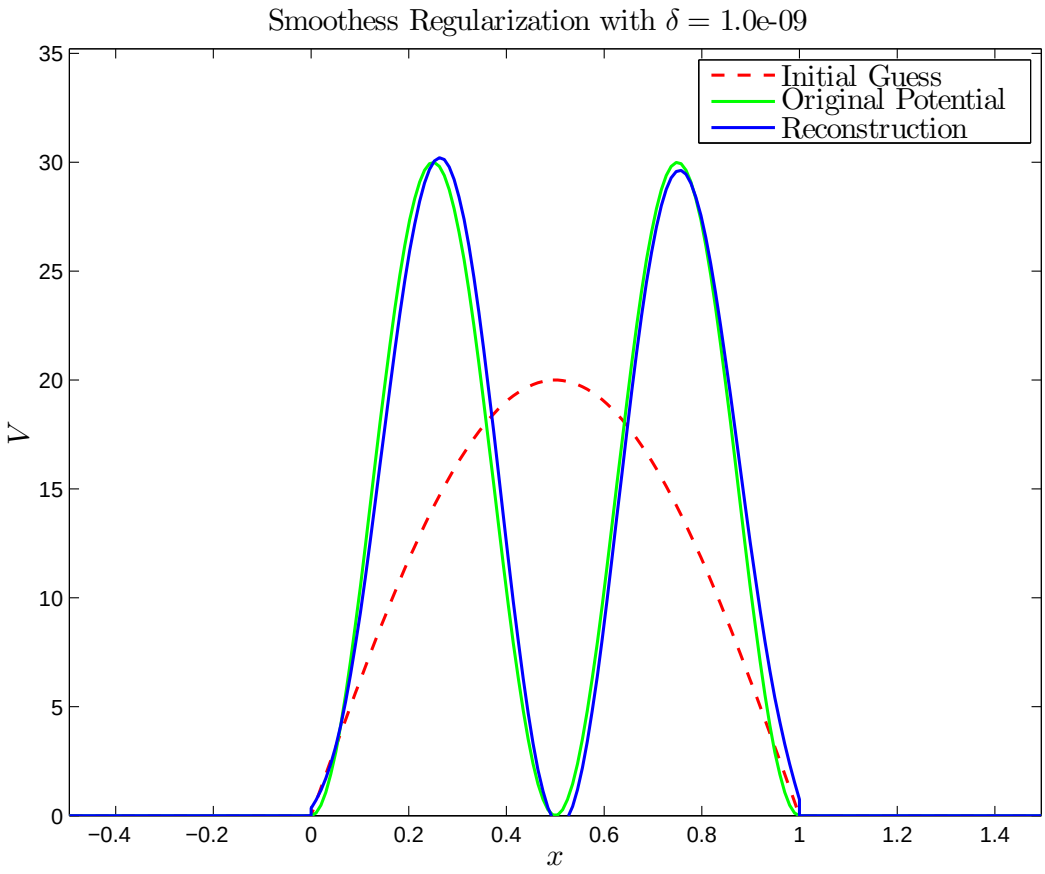

Figure 4.4: Reconstruction of a double-bump potential barrier.

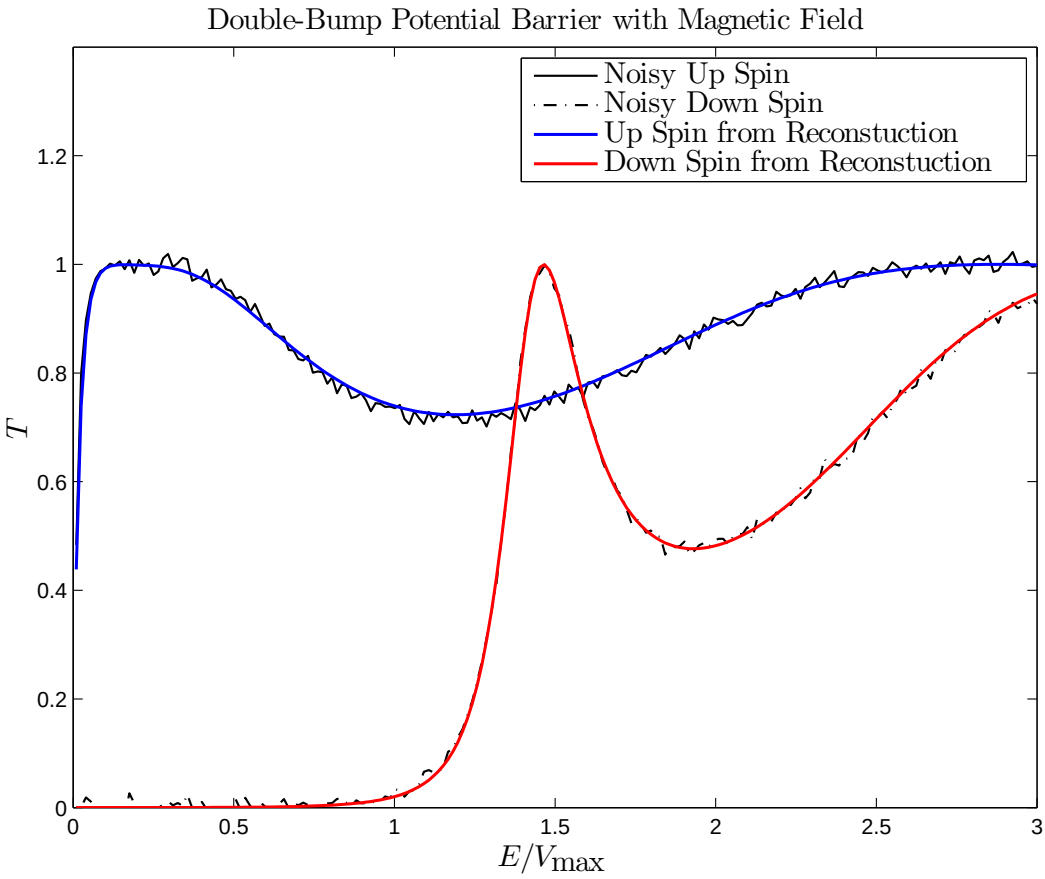

Figure 4.5: Transmission probability of the reconstructed double-bump potential barrier.

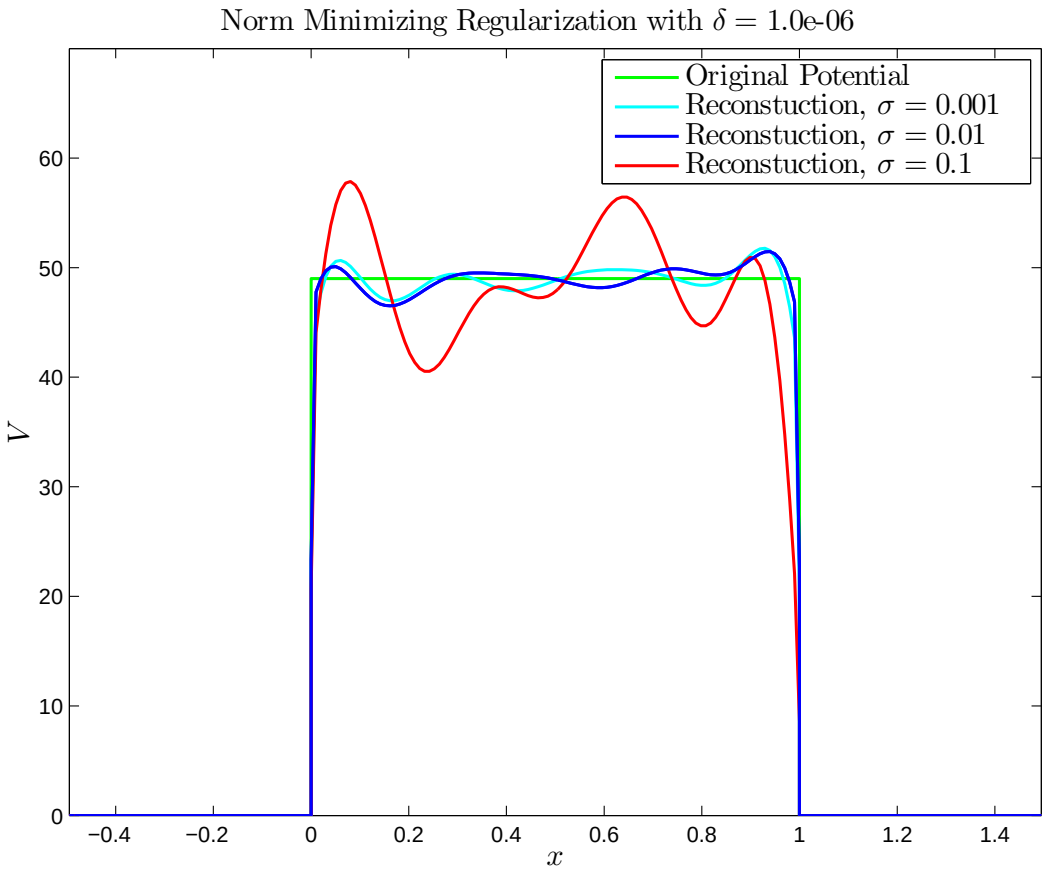

Figure 4.6: Noise comparison for reconstructions of a constant potential barrier.

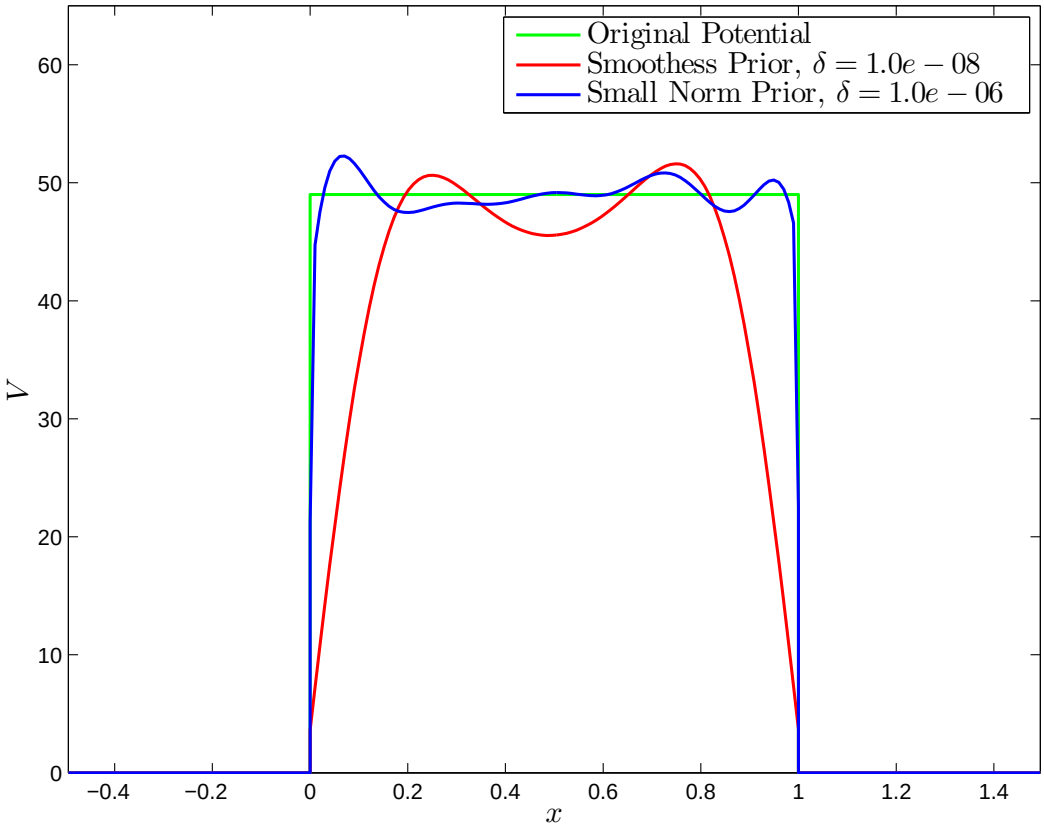

Figure 4.7: Regularization comparison for reconstructions of a constant potential barrier.

Chapter 5

Conclusions

In this work we have presented the formulation of one-dimensional quantum scattering, and provided an algorithm for computing solutions to the forward and inverse problems. The considered model included both electric potential and magnetic field. The Schrödinger equation for this setup was derived in Chapter 2. In Chapter 3, we showed that the forward scattering problem is uniquely solvable and proceeded by presenting a finite element solver for the corresponding boundary value problem. The sensitivity of the solver with respect to the input data, the electric potential, was investigated in Chapter 3. Information on the sensitivity is important when inverting the scattering problem: the non-linear dependence between the solution and the initial data prevents an application of straightforward inverse algorithms designed for linear problems. In Chapter 4, we described a sensitivity-based iterative scheme for reconstructing the potential from the known scattering probabilities. We also presented numerical examples of reconstructed potentials with different forms of regularization. Regularization had a significant effect on the reconstructions: with a proper choice, good results were achieved but even small changes in the amount of regularization lead to erratic behavior.

The magnetic field was included in the forward problem to obtain physically more interesting model that contains spins of electrons and multiple transmission coefficients. Reconstructing the magnetic field in addition to the potential function was excluded from the inverse algorithm. Such considerations could be included in future studies. The inclusion of the magnetic field adds more complexity because one would have to reconstruct four sets of parameters, one for the potential and three for the field, from four transmission probabilities. Even though there would be as 'much' data as parameters, difficulties could arise from the transmission probabilities having underlying interdependencies.

In Chapter 2 we briefly discussed a slightly different approach to formal-

izing the scattering problem. Instead of considering only one incoming wave, we assumed two waves meeting at the scattering region and then rebounding back or transmitting through the scatterer. This setup is called the S-matrix formalism; the S-matrix describes the transformation of the input into the output. This enables studying cases where the incoming waves can enter from multiple directions. Updating the inversion method to support the S-matrix formalism would allow easier examination of multi-dimensional models in the future.

We could also regard a different aspect of the problem: including additional constraints in the algorithm and turning it effectively into an optimization procedure rather than a sole inverse solver. This way it would be possible to pose questions, such as, is it possible to control the transmission of both spin components in a specified energy interval? Or, how the maximal total transmission can be achieved at a certain energy? If we knew that the incoming electrons have some energy distribution, would it be possible to create adjustable spin filter by only controlling the potential, the magnetic field or both?

Bibliography

[1] Andrews, B., and Hopper, C. *The Ricci Flow in Riemannian Geometry. A Complete Proof of the Differentiable 1/4-Pinching Sphere Theorem.* Springer, 2011.

[2] Bruus, H., and Flensberg, K. *Many-body quantum theory in condensed matter physics*, 2003 ed. Ørsted Laboratory, Niels Bohr Institute, University of Copenhagen, and Mikroelektronik Centret, Technical University of Denmark, 2003.

[3] Cakoni, F., and Colton, D. *A qualitative approach to inverse scattering theory.* Springer, 2014.

[4] Coddington, E. A., and Levinson, N. *Theory of ordinary differential equations.* McGraw-Hill, 1955.

[5] Colton, D., and Kress, R. *Inverse acoustic and electromagnetic scattering theory*, third ed. Springer, 2013.

[6] Datta, S. *Electronic transport in mesoscopic systems.* Cambridge university press, 1997.

[7] Dautray, R., and Lions, J.-L. *Mathematical analysis and numerical methods for science and technology. Vol. 2. Functional and variational methods.* Springer-Verlag, 1988.

[8] Dierkes, T., Dorn, O., Natterer, F., Palamodov, V., and Sielschott, H. Frechet derivatives for some bilinear inverse problems. *SIAM J. Appl. Math. 62* (2002), 2092–2113.

[9] Engl, H. W., Hanke, M., and Neubauer, A. *Regularization of inverse problems.* Kluwer, 1996.

[10] Esterhazy, S., and Melenk, J. M. On stability of discretizations of the helmholtz equation. In *Numerical analysis of multiscale problems*, Lect. Notes Comput. Sci. Eng., 83. Springer, 2012, pp. 285–324.

[11] Kaipio, J. P., and Somersalo, E. *Statistical and Computational Inverse Problems*, first ed. Spinger, 2005.

[12] Lax, P. D., and Phillips, R. S. *Scattering theory*, second ed. Academic Press, Inc., 1989.

[13] Liboff, R. L. *Introductory Quantum Mechanics*, fourth ed. Addison Wesley, 2003.

[14] Nocedal, J., and Wright, S. J. *Numerical Optimization*, first ed. Spinger, 1999.